

## Article

# Did You Ask for Citations? An Insight into Preprint Citations *en route* to Open Science

Mario Pagliaro 

Istituto per lo Studio dei Materiali Nanostrutturati, CNR, via U. La Malfa 153, 90146 Palermo, Italy;  
mario.pagliaro@cnr.it

**Abstract:** This study investigates citation patterns between 2017 and 2020 for preprints published in three preprint servers, one specializing in biology (bioRxiv), one in chemistry (ChemRxiv), and another hosting preprints in all disciplines (Research Square). Showing evidence that preprints are now regularly cited in peer reviewed journal articles, books, and conference papers, the outcomes of this investigation further substantiate the value of open science also in relation to citation-based metrics on which the evaluation of scholarship continues to rely on. This analysis will be useful to inform new research-based education in today's scholarly communication.

**Keywords:** preprint; citations; scholarly communication; open science; peer review; impact factor



**Citation:** Pagliaro, M. Did You Ask for Citations? An Insight into Preprint Citations *en route* to Open Science. *Publications* **2021**, *9*, 26.  
<https://doi.org/10.3390/publications9030026>

Academic Editor: Teja Koler Povh

Received: 15 April 2021

Accepted: 21 June 2021

Published: 25 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Following the introduction of the world wide web in 1989, it became clear that the internet could reshape scholarly communication. Research papers would no longer be printed in paper journals, but would rather be published in digital format on the web after peer review, or even before peer review in the form of “preprints”. Particle (“high-energy” in physics jargon) physicists used to circulate paper-copy preprints since 1969, when a database management system aggregating preprints shared between different institutions was established in Stanford, an urban area in California hosting a large particle accelerator. In 1991, the Stanford Physics Information Retrieval System became the first database accessible through the web [1].

In 2003, Harnad forecast that “once the preprints and postprints of all 2 million articles appearing annually in the world’s 20,000 peer-reviewed journals are openly accessible, research progress will become much more rapid and interactive” [2]. Two decades later, the number of preprints published in 2019 nearly reached the 227,000 threshold [3]. Regardless of the exponential growth (doubling of preprint number in less than 10 years) observed since 1991, the ratio of preprints to all scientific articles published in the first nine months of 2020 was just 6.4% [3]. Only two basic science disciplines have shown a significant uptake of preprints, namely: physics (close to 35% preprints to all-papers published in 2019) and mathematics (about 30% in 2019) [3].

Hosting close to 1,800,000 preprints by late 2020, the world’s largest preprint server managed by the Library of Cornell University at <https://arxiv.org> (accessed on 15 April 2021), in 2020, reached a 14,861/month publication rate [4]. For comparison, the main biology preprint server managed by the USA-based Cold Spring Harbor Laboratory at <https://biorxiv.org> (on 15 April 2021), in 2020, reached a publication rate of 3193/month [5].

From Advance to Zenodo, as of April 2021, a list of 68 preprint servers were found in the directory made openly accessible on the internet by Rittman [6], a manager at CrossRef, a not-for-profit organization for scholarly publishing. Another comprehensive list of preprint servers in early 2020 found 57 preprint servers [7]. These repositories host preprints either in specific research fields (EarthArXiv, bioRxiv, etc.) or in all scientific disciplines (SSRN, Preprints, Authorea, etc.).

According to a reputed manual for authors of research papers written in English published in 2017, “not having been subject to peer review preprints are treated as unpublished material” [8]. Fifteen years before, between 2002 and July 2003, renowned mathematician Perelman published the first of three papers, which appeared only in preprint form, in which he provided proof of the geometrization conjecture, including the Poincaré conjecture. By early April 2021, the first preprint study [9] had been cited 2503 times [10]. This fact alone shows evidence that the scientific community has never seen preprints as “unpublished material”, but rather as scientific articles that are read, studied, and cited, even though not having gone through the peer review process.

In brief, preprints are regularly cited in peer reviewed journals, books, conference papers, and presentations. For example, by late 2016, the preprints posted in arXiv between its launch on August 1991 and 2016 had received 135,782 citations, of which 23,288 were from 2016 [11]. Previous studies have investigated citations for preprints in arXiv [11–14]. This study investigates citation patterns between 2017 and 2020 for preprints published in two specialized preprint servers publishing research articles in chemistry (ChemRxiv) and in biology (bioRxiv), and between 2018 and 2020 in one multidisciplinary preprint repository (Research Square). As the evaluation of scholarship continues to rely on citation-based metrics [15], this analysis and its outcomes will be useful to inform new research-based education in today’s scholarly communication [16].

## 2. Methods

An online search for the total citations of the research papers published in ChemRxiv and bioRxiv between 2017 and 2020 was conducted in Scopus on 7 April 2021. A similar search was carried out in Dimensions for papers posted in Research Square between 2018 and 2020. The publishing dates researched were not the same because the Research Square preprint server made its debut in October 2018.

Launched in 2004 by Elsevier, Scopus (<https://www.scopus.com>, accessed on 15 April 2021) is a scholarly database indexing scientific journals, books, and conference proceedings that, by late 2019, included 23,452 active journal titles, about 120,000 conferences, and 206,000 books from more than 5000 publishers, adding some 3 million records every year [17]. Launched in 2018, Dimensions (<https://www.dimensions.ai>, accessed on 15 April 2021) is the most comprehensive bibliographics database, with 109 million publications indexed and about 1.1 billion citations as of September 2019. As of May 2020, Dimensions indexed more than 74,000 journals [17]. Data on cumulative preprints were obtained from different sources, namely: [18] for bioRxiv, [19] for ChemRxiv, and [20] for Research Square.

## 3. Results

The results in Table 1 show the large difference in the uptake of preprints between chemists and biologists. By the end of 2020 bioRxiv, launched in November 2013, hosted 107,518 preprints, whereas ChemRxiv, launched on August 2017, hosted 6127 preprints. Among scholars in the so-called “basic” or “exact” sciences, indeed, chemists are those with the lowest uptake of preprints [21] as well as of the open access (OA) [22] publishing models.

In one year only, Research Square went from 5666 preprints hosted in 2019, to 38,448 in 2020 (+580%). Since then, as of mid April 2021, the website added another 32,251 preprints, thus approaching the 80,000 (77,641) threshold [23]. Growth was chiefly driven by preprints dealing with COVID-19, which led Research Square to rapidly become one of the top three preprint servers by volume for research related to the disease [24]. Researchers likely opted to publish findings and reviews on the disease via a platform whose majority owner (Springer Nature) publishes some of the leading medicine journals. However, the platform publishes preprints in all scientific disciplines, including language studies, philosophy, and law, with submissions rapidly increasing in many other fields. By April 2021, for example, the server hosted 636 preprints in the chemical sciences.

The list of the top five journals citing preprints in bioRxiv and ChemRxiv (Table 2) shows that the publications most frequently citing preprints posted in these repositories are reputed multidisciplinary or specialized chemistry and biology journals.

**Table 1.** Preprint citations and number of preprints of selected preprint servers for 2017–2020.

Preprint Server	Year	Citations <sup>a,b</sup>	Cumulative Preprints <sup>c</sup>
bioRxiv	2020	23,820	107,518
	2019	11,280	68,801
	2018	5880	39,620
	2017	2643	18,837
ChemRxiv	2020	1432	6127
	2019	437	2289
	2018	85	1053
	2017	7	350
Research Square	2020	3208	38,448
	2019	45	5666
	2018	8	2
	2017	-	-

<sup>a</sup> Source: Scopus, 2021; <sup>b</sup> Source: Dimensions, 2021; <sup>c</sup> Source: Refs. [18–20].

**Table 2.** Top five journals citing preprints in bioRxiv and ChemRxiv, and type of citing documents <sup>a</sup>.

Rank	Journal	Citations	Document Type	Documents Citing (Share)
bioRxiv				
1	<i>Scientific Reports</i>	1116	Article	37,694 (70%)
2	<i>eLife</i>	954	Review	12,567 (23.5%)
3	<i>Plos ONE</i>	923	Conference paper	1730 (3.2%)
4	<i>Nature Communications</i>	787	Book chapter	939 (1.75%)
5	<i>International Journal of Molecular Sciences</i>	685	Note	584 (1.1%)
ChemRxiv				
1	<i>Organic Letters</i>	90	Article	1788 (71.6%)
2	<i>Journal of the American Chemical Society</i>	70	Review	616 (24.7%)
3	<i>Angewandte Chemie International Edition</i>	66	Book chapter	35 (1.4%)
4	<i>Journal of Biomolecular Structure and Dynamics</i>	45	Conference paper	33 (1.3%)
5	<i>Journal of Chemical Information and Modeling</i>	44	Note	26 (1%)

<sup>a</sup> Source: Scopus, 2021.

Preprints in bioRxiv and ChemRxiv are chiefly cited in journal original research and review articles, in both cases accounting for >93% of the citations. Compared with preprints posted at ChemRxiv, preprints in bioRxiv are cited three times more frequently in conference papers.

Contrary to the preprints posted in ChemRxiv, preprints posted in bioRxiv and Research Square are chiefly cited by OA journals. This is not surprising, considering that chemists are the researchers with the lowest fraction of published OA papers [21]. On the other hand, the top three journals citing ChemRxiv preprints (*Organic Letters*, *Journal of the American Chemical Society*, and *Angewandte Chemie International Edition*) are among the most reputed chemistry journals. The remaining two (*Journal of Biomolecular Structure and Dynamics* and *Journal of Chemical Information and Modeling*) are well known structural chemistry journals.

It is also relevant that the second subject area citing preprints in ChemRxiv (Table 3) is “biochemistry, genetics, and molecular biology”. In the course of 2020, the latter became

the main subject area of preprints posted in bioRxiv. Until late 2018, the main subject area for preprints posted on bioRxiv was neuroscience, which, by October 2018, had become the first bioRxiv collection to cross the 6000 preprint threshold [18].

**Table 3.** Top five subject areas citing preprints in bioRxiv and ChemRxiv <sup>a</sup>.

Rank	Subject Area	Preprints
<b>bioRxiv</b>		
1	Biochemistry, genetics, and molecular biology	26,488
2	Medicine	16,875
3	Agricultural and biological sciences	11,021
4	Immunology and microbiology	8119
5	Neuroscience	7710
<b>ChemRxiv</b>		
1	Chemistry	1421
2	Biochemistry, genetics, and molecular biology	780
3	Chemical engineering	595
4	Materials science	438
5	Medicine	326

<sup>a</sup> Source: Scopus, 2021.

These findings also confirm that in the case of chemistry and biology, preprints reach a much broader readership. Preprints in bioRxiv are vastly cited, for instance, by researchers in agricultural sciences, whereas preprints in ChemRxiv are twice more frequently cited by researchers in biochemistry than by scientists working in materials science (Table 3).

Table 4 shows the top 10 research fields and number of citations for preprints in Research Square as of early April 2021. Entry 9 shows evidence that preprints in the field of medical microbiology posted in Research Square are cited at a significantly higher rate when compared, for instance, to preprints in neurosciences (entry 10 in Table 4) [20].

**Table 4.** Top 10 research fields and number of citations for preprints in Research Square as of 7 April 2021.

Entry	Research Category	Preprints	Citations	Citation/Preprint
1	Medical and health sciences	33,396	3546	0.11
2	Clinical sciences	12,024	1062	0.09
3	Public health and health services	11,930	1345	0.11
4	Biological sciences	7123	690	0.10
5	Oncology and carcinogenesis	4740	130	0.03
6	Genetics	3941	233	0.06
7	Biochemistry and cell biology	2375	397	0.17
8	Cardiorespiratory medicine and haematology	1880	153	0.08
9	Medical microbiology	1653	681	0.41
10	Neurosciences	1499	95	0.06

There are no clear trends. For example, by early April 2021, the 101 preprints in analytical chemistry (not shown in Table 4) received 20 citations, with a high (0.2) citation/preprint ratio, whereas by early April 2021, the 37 preprints in nanotechnology on Research Square had never been cited. In general, between 2019 and 2020, the growth in citation rate for preprints in Research Square (+7000%) was one order of magnitude higher than the increase in the number of preprints published (+580%). In detail, citations went from 45 (for 5666 preprints) in 2019 to 3208 (for 38,848 preprints) in 2020 [20].

#### 4. Discussion

A citation to a preprint server is a citation that will not be counted by the scholarly database used by the commercial USA-based company (Clarivate Analytics) that publishes the journal impact factor (JIF) of indexed journals every year. This may explain the “con-

cerns” found in the 2018 report of the International Scientific Publishers Association, that, “preprints (which can be brought up to date) may become a go-to place for the version of record, undermining publisher business models”, leading to “concerns over the loss of citations from journals to preprints servers, with well over 8000 citations to bioRxiv reported on Web of Science [25]”.

As the outcome of a highly skewed distribution, for which typically 15% of the papers in a journal account for half the total citations [26], the impact factor is a misleading statistical indicator because the vast majority (typically 85%) of the journal’s articles will actually get fewer citations than indicated by the JIF. Because of this, namely that the JIF is a statistically poor indicator, it should not be used to evaluate research [27]. As noted by Curry in 2012, “if you are judging grant or promotion applications and find yourself scanning the applicant’s publications, checking off the impact factors, you are statistically illiterate” [28]. Furthermore, empirical evidence derived from analyzing measures of experimental and statistical rigor concerning data reported and the use of statistics in neuroscience and psychology, reveals that “elite” (high JIF) scientific journals are those with the lowest reliability of published research [29].

Investigating the fundamental cause of the ongoing “impact factor mania”, Casavedell and Fang suggested that “the impact factor mania persists because it confers significant benefits to individual scientists and journals. Impact factor mania is a variation of the economic theory known as the ‘tragedy of the commons,’ in which scientists act rationally in their own self-interests despite the detrimental consequences of their actions on the overall scientific enterprise [30]”.

This, in turn, explains why “elite” journals do not expand to accommodate all meritorious articles, continuing in what Fang and Casavedell call “artificial restrictions on journal size serve to perpetuate the current wasteful system that requires authors to cascade serial submissions from one journal to another” [30]. Put simply, by expanding the journal size beyond a certain threshold, the denominator in the equation affording the JIF would rapidly become too large, and the JIF would decrease, regardless of the increase in citations in the JIF equation’s numerator. This is what happened, for example, in the case of so-called “mega journals” (large-volume OA journals also accepting replication studies and negative results), whose JIFs rapidly decreased a few years after their inception [31].

As noted by Davis, a citation “is the basis for a system of rewarding those who make significant contributions to public science” [32]. This is the original idea that led Garfield to introduce the JIF in 1955 [33]—scientific quality is associated with citations from peers. Hence, there is no logical basis for promotion and tenure committees or for research funding agencies to evaluate a citation differently to a preprint or to a peer reviewed paper. Accordingly, in the USA, in early 2017, the National Institutes of Health recommended applicants and awardees to include in their applications, proposals, and reports the citations of any “interim research product” [34]. In other words, along with enhanced visibility and professional and funding opportunities [35], the benefits of preprints include a higher number of citations directly in terms of preprint citations. If the preprint is subsequently published in a peer reviewed journal, the journal article (also called “version of record”) can be merged with the preprint, allowing the database to add the citations (easily done for instance on Google Scholar by selecting the preprint and the journal article and choosing “Merge” from the “Actions” menu).

Studying 1495 mathematics cited preprints posted in arXiv, scientometrics scholars have lately unveiled three important findings, namely: (i) 71.8% of cited preprints are cited before journal publication; (ii) about 50% of all preprint citations are received within 24 months of publication, after which the average citation rate of preprints decreases because of journal publication (authors prefer to cite the journal version rather than the preprint when both are available); and (iii) the preprint version and the journal version of the same papers have different readerships, with the preprint server reaching a much broader readership, as shown by the fact that 27.5% of the total preprint citations (vs.



12.5% for the journal versions) originate from papers assigned to another discipline beyond mathematics [13].

## 5. Teaching Open Science

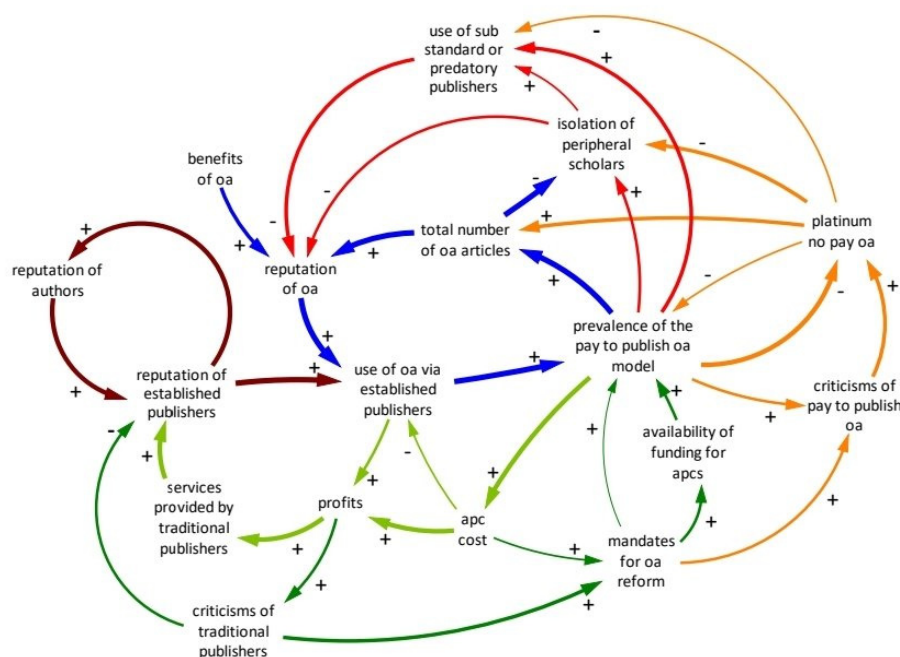
When teaching open science, to win skepticism of young researchers working under the “publish-or-perish” pressure [36], I suggest the use of practical examples as case studies. For instance, in 2018, Kievit and co-workers published a preprint on raincloud plots (a new data visualization tool providing bias-free statistical information) in PeerJ Preprints [37]. In the second version of the preprint posted in Wellcome Open Research eight months later, the team inserted a new section illustrating the benefits derived from posting a preprint, which is worth reading; “Firstly, posting the manuscript as preprint has vastly widened the reach. To date (March 2019) our preprint was viewed 9803 times, with 6309 downloads. However, views and downloads alone don’t necessarily entail engagement. Since publication the preprint alone has already been cited 18 times. Moreover, in depth engagement has gone well beyond mere citations. Several individuals have created their own useful tutorials, summarizing our paper and asking useful questions, posted constructive criticism, discussed raincloud plots as part of various plotting alternatives, created a shiny app, wrote an accessible tutorial using native *R* datasets, a new package, creating various animated interactive visualizations. Our codebase itself received feedback through various avenues including formal pull requests on github, comments on the preprint, twitter replies and email. In this new version of our paper we have tried our best to integrate all these suggestions and comments, which without fail have improved the usability of our code”. In addition, “Social media, specifically twitter, provided the central hub where all these benefits coalesced. The paper has been tweeted at least 750 times, with an estimated reach of up to 1,500,000 total followers, and as such is the principal driver for the engagement our preprint has received. This engagement has yielded invaluable feedback, comments, and suggestions, and were even lucky enough to track down the first instance of an early precursor of the raincloud plot (Ellison, 2018). Moreover, the paper itself was inspired by a twitter discussion, and brings together co-authors who have never met in person. Together, these interactions illustrate the fundamentally two-way street of new publishing models, which facilitate access without paywalls and allow for near instantaneous improvements to ongoing work [38].”

Making “near instantaneous improvements to ongoing work” possible, preprints allow for reaping the benefits of open science, at the same time enhancing the number of citations and related citation-based metrics that continue to be used for appraising scientists [15,39].

It is enough to review the citations of preprints posted in arXiv to realize that physicists and mathematicians have never made a distinction between the citation of a study deposited in arXiv and the citation of a study in a peer reviewed journal. In other words, they never added the “preprint” word in the reference, as required by certain research funding agencies, such as the aforementioned National Institutes of Health, and by certain journals. Scholars in physics know, for instance, that reputed journals such as *Physical Review*, until 1960, used peer review for only half of the papers received, and even in that case, peer review consisted of the editor asking one referee an opinion on a manuscript for which the editor needed advice [40]. From Krebs’ 1937 work on the citric acid cycle rejected by *Nature* and published in *Experientia*, to Ernst’s 1966 work on nuclear magnetic resonance spectroscopy rejected twice by the *Journal of Chemical Physics* to be finally published in the *Review of Scientific Instruments*, through Mullis’ discovery of the polymerase chain reaction rejected by *Science* and published in *Methods in Enzymology*, the fact that several discoveries leading to major scientific progress (and to Noble Prizes) were actually rejected by peer reviewed journals shows evidence that the peer review system has significantly delayed scientific progress, and perhaps also suppressed it [41].

To avoid these negative effects, scholars today can publish their findings in preprint form immediately, and then seek publication of the preprint in OA or paywalled peer

reviewed journals. Finding that nearly 50% of all research papers published in 2020 were open accessible, Dudley recently unveiled the main drivers and obstacles to open access publishing [42]. The dominance of the pay-to-publish OA model and its reinforcing factors led to mandates for OA publishing, which continued to focus on the pay-to-publish model (Figure 1).



**Figure 1.** Dudley's closed loop diagram describing the current (early 2021) situation in the scientific publishing system. Arrows (causal links), indicate how a change in the causal variable affects change in the second variable. Change in the same direction is indicated with a plus sign. Change in the opposite direction is indicated with a minus sign. (Reproduced from [42], Creative Commons CC BY 4.0 license).

The financial barrier created by the latter model in turn created a demand for sub-standard journals from authors in low income countries, reinforcing the dichotomy between scholars in wealthier and those in poorer countries. Indeed, in 2017, the largest number of researchers who published in “predatory” journals were from India, Nigeria, Turkey, and other economically developing countries [43]. The outcome of this situation is that the article processing charge (APC) of “elite” OA journals currently exceeds \$3000 and at times even the \$5000 thresholds (Table 5), while in poorer countries, numerous sub-standard OA journals have emerged charging low APCs.

As mentioned above, to escape this vicious circle, scholars today may first freely publish the outcomes of their work in preprint form, and then submit the preprint for publication after peer review in low cost (or in free) OA journals, or even in paywalled journals, taking care however to “green” self-archive their papers [2]. On the other hand, Laakso has found that in 2010, a share exceeding 80% of all articles published that year in the studied fields (almost 1 million out of 1.1 million articles investigated) could be self-archived after 12 months of publication [44]. Yet, only 12% of said articles were found to be actually self-archived, with scholars in certain disciplines such as chemistry and chemical engineering not even reaching 10% [45]. Nearly ten years later, extending this analysis to all articles indexed in the Web of Science (a commercial bibliographic database), the share of green self-archived articles was found to be 4%, with another 7% made OA directly by journal editors [46]. This, once again, reveals the widespread need for scholars in all disciplines to have updated education of the practical value in the field of open science [16].

**Table 5.** Article processing charge (APC) for selected elite scientific journals as of April 2021.

Journal	Publisher	APC (USD)
<i>Nature Communications</i>	Springer Nature	5560 <sup>a</sup>
<i>Advanced Science</i>	Wiley	5000 <sup>b</sup>
<i>JACS Au</i>	American Chemical Society Publishing	5000 <sup>c</sup>
<i>Science Advances</i>	American Association for the Advancement of Science	4500 <sup>d</sup>
<i>PLOS Biology</i>	Public Library of Science	4000 <sup>e</sup>
<i>eLife</i>	eLife Sciences Publications	3000 <sup>f</sup>
<i>Cell Reports</i>	Elsevier	5200 <sup>g</sup>

<sup>a</sup> (The Americas, Greater China, and Japan), <https://www.nature.com/ncomms/about/article-processing-charges> (accessed on 15 April 2021). <sup>b</sup> <https://authorservices.wiley.com/asset/Wiley-Journal-APCs-Open-Access.xlsx> (accessed on 15 April 2021). <sup>c</sup> In case of (CC-BY) license, in case of (CC-BY-NC-ND) license, \$4000, <https://acsopencscience.org/open-access/pricing/> (accessed on 15 April 2021). <sup>d</sup> <https://advances.sciencemag.org/content/licensing-and-charges>. <sup>e</sup> <https://plos.org/publish/fees/> (accessed on 15 April 2021). <sup>f</sup> <https://reviewer.elifesciences.org/author-guide/fees> (accessed on 15 April 2021). <sup>g</sup> <https://www.cell.com/rights-sharing-embargoes> (accessed on 15 April 2021).

On the other hand, even though scholars can self-archive their papers after the embargo period, the two main reasons to publish their research in preprint form remain unchanged, namely: (i) immediate communication of research findings, and (ii) unlimited, free access to said findings from the scholarly community thanks to the OA nature of all preprint servers, so that authors can get immediate feedback [16,36]; or, as suggested by an anonymous reviewer of the present study, “the preprint breaks the time difference between regions through social media, so that even if you are not a fellow scholar, you can give timely opinions from the perspective of an observer. It increases objectivity and saves time. It is far better than the publication system of journal articles that delays scientific progress [47]”.

## 6. Conclusions

Showing evidence that preprints are now regularly cited in peer reviewed journal articles, books, and conference papers by investigating the citation patterns for preprints published in ChemRxiv, Research Square, and bioRxiv, this study further substantiates the value of open science in relation to citation-based metrics on which the evaluation of scholarship continues to rely on [39]. In brief, gone are the days in which preprints, for example those published in the early years of bioRxiv, were highly read and shared online, but poorly cited [48].

Driven by the scholarship evaluation system, largely based on citations and achievements in securing research funds, researchers will continue to publish in peer reviewed journals. To benefit from the unique advantages that the practice of open science offers to their own career in terms of enhanced citations and visibility, the very same scholars will inevitably adopt preprints and self-archiving. Accordingly, the number of journals that do not accept the submission of preprints already posted on an open access server has now nearly vanished. Virtually all major publishers (Elsevier, Springer Nature, Wiley, MDPI, Informa, and Oxford University Press, as well as all other 14 publishers comprising the top 20 world’s publishing companies [49]) today “encourage posting of preprints of primary research manuscripts” [50]. Many of them actually purchased existing preprint servers or launched their own preprint repositories. MDPI, for instance, owns Preprints, Elsevier bought SSRN, Wiley acquired Authorea, and Springer Nature bought Research Square.

In the open science era, purposeful evaluation of scholarship includes the evaluation of preprints [39], and thus takes into account the number of preprint citations. In conclusion, along with the list of preprints and the number of preprint citations, I recommend scholars also include in their CV the value of alternative metrics indicators (such as those provided by Altmetric, Mendeley, and PlumX) which track and report alternative metrics data measuring the online impact of research articles [51].



**Funding:** This research received no external funding.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Acknowledgments:** This study is dedicated to Maria Pia Virzì for many years of friendship.

**Conflicts of Interest:** The author declares no conflict of interest.

## References and Note

- Gentil-Beccot, A.; Mele, S.; Brooks, T.C. Citing and reading behaviours in high energy physics: How a community stopped worrying about journals and learned to love repositories. *arXiv* **2009**, arXiv:0906.5418. Available online: <http://arxiv.org/ftp/arxiv/papers/0906/0906.5418.pdf> (accessed on 15 April 2021).
- Harnad, S. Electronic preprints and postprints. In *Encyclopedia of Library and Information Science*; Marcel Dekker: New York, NY, USA, 2003; Available online: <http://www.ecs.soton.ac.uk/~harnad/Temp/eprints.htm> (accessed on 15 April 2021).
- Xie, B.; Shen, Z.; Wang, K. Is preprint the future of science? A thirty year journey of online preprint services. *arXiv* **2021**, arXiv:2102.09066. Available online: <https://arxiv.org/abs/2102.09066> (accessed on 15 April 2021).
- Cornell University. arXiv Submission Rate Statistics. 2021. Available online: [https://arxiv.org/help/stats/2020\\_by\\_area/index](https://arxiv.org/help/stats/2020_by_area/index) (accessed on 6 April 2021).
- Rxivist. Site-Wide Metrics. 2021. Available online: <https://rxivist.org/stats> (accessed on 6 April 2021).
- Rittman, M. Research Preprints: Server List. 2021. Available online: <https://docs.google.com/spreadsheets/d/17RgfuQcGJHKSsSJwZZn0oiXAnimZu2sZsWp8Z6ZaYYo/> (accessed on 6 April 2021).
- ASAPBio. Directory of Preprint Server Policies and Practices. 2020. Available online: <https://asapbio.org/preprint-servers> (accessed on 6 April 2021).
- The Chicago Manual of Style*, 17th ed.; Section 14.173: Journal article preprints; The University of Chicago Press: Chicago, IL, USA, 2017.
- Perelman, G. The entropy formula for the Ricci flow and its geometric applications. *arXiv* **2002**, arXiv:math/0211159. Available online: <https://arxiv.org/abs/math/0211159> (accessed on 15 April 2021).
- Google. Google Scholar. Available online: <http://scholar.google.com> (accessed on 15 April 2021).
- Noruzi, A. ArXiv popularity from a citation analysis point of view. *Webology* **2016**, *13*, 22. Available online: <http://www.webology.org/2016/v13n2/editorial22.pdf> (accessed on 15 April 2021).
- Ferrer-Sapena, A.; Aleixandre-Benavent, R.; Peset, F.; Sanchez-Perez, E.A. Citations to arXiv preprints by indexed journals and their impact on research evaluation. *J. Inf. Sci. Theory Pract.* **2018**, *6*, 6–16. [CrossRef]
- Wang, Z.; Chen, Y.; Glänzel, W. Preprints as accelerator of scholarly communication: An empirical analysis in Mathematics. *J. Inf.* **2020**, *14*, 101097. [CrossRef]
- Traag, V. Inferring the causal effect of journals on citations. *Quant. Sci. Stud.* **2021**, 1–9. [CrossRef]
- Ioannidis, J.P.; Boyack, K.W. Citation metrics for appraising scientists: Misuse, gaming and proper use. *Med. J. Aust.* **2020**, *212*, 247–249.e1. [CrossRef]
- Pagliaro, M. Publishing scientific articles in the digital era. *Open Sci. J.* **2020**, *5*, 3. [CrossRef]
- Singh, V.K.; Singh, P.; Karmakar, M.; Leta, J.; Mayr, P. The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometrics* **2021**, *126*, 5113–5142. [CrossRef]
- Available online: [http://api.biorxiv.org/reports/content\\_summary](http://api.biorxiv.org/reports/content_summary) (accessed on 7 April 2021).
- GitHub. ChemRxiv Dashboard. 2021. Available online: <https://chemrxiv-dashboard.github.io> (accessed on 7 April 2021).
- Available online: <https://app.dimensions.ai/> (accessed on 7 April 2021).
- Piwowar, H.; Priem, J.; Larivière, V.; Alperin, J.P.; Matthias, L.; Norlander, B.; Farley, A.; West, J.; Haustein, S. The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ* **2018**, *6*, e4375. [CrossRef] [PubMed]
- Pagliaro, M. Preprints in chemistry: An exploratory analysis of differences with journal articles. *Publications* **2021**, *9*, 5. [CrossRef]
- Available online: <https://www.researchsquare.com/browse> (accessed on 8 April 2021).
- Else, H. How a torrent of COVID science changed research publishing—In seven charts. *Nat. Cell Biol.* **2020**, *588*, 553. [CrossRef]
- The STM Report*, 5th ed.; International Association of Scientific, Technical and Medical Publishers: The Hague, The Netherlands, 2018; p. 10.
- Seglen, P.O. The skewness of science. *J. Assoc. Inf. Sci. Technol.* **1992**, *43*, 628–638. [CrossRef]
- Seglen, P.O. Why the impact factor of journals should not be used for evaluating research. *BMJ* **1997**, *314*, 497. [CrossRef] [PubMed]
- Curry, S. Sick of Impact Factors. 13 August 2012. Available online: <http://occamstypewriter.org/scurry/2012/08/13/sick-of-impact-factors> (accessed on 7 April 2021).
- Brembs, B. Prestigious Science Journals Struggle to Reach Even Average Reliability. *Front. Hum. Neurosci.* **2018**, *12*, 37. [CrossRef] [PubMed]
- Casadevall, A.; Fang, F.C. Causes for the Persistence of Impact Factor Mania. *mBio* **2014**, *5*, e00064-14. [CrossRef]
- Brainard, J. Open-access megajournals lose momentum as the publishing model matures. *Science* **2019**. [CrossRef]
- Davis, P. Journals Lose Citations to Preprint Servers. 2018. Available online: <https://scholarlykitchen.sspnet.org/2018/05/21/journals-lose-citations-preprint-servers-repositories> (accessed on 15 April 2021).

33. Garfield, E. The History and Meaning of the Journal Impact Factor. *JAMA* **2006**, *295*, 90–93. [CrossRef]
34. National Institutes of Health. Reporting Preprints and Other Interim Research Products, Notice Number: NOT-OD-17-050. 24 March 2017. Available online: <https://grants.nih.gov/grants/guide/notice-files/not-od-17-050.html> (accessed on 7 April 2021).
35. Puebla, I.; Polka, J.; Rieger, O. Preprints: Their evolving role in science communication. *MetaArXiv* **2021**. [CrossRef]
36. Tian, M.; Su, Y.; Ru, X. Perish or Publish in China: Pressures on Young Chinese Scholars to Publish in Internationally Indexed Journals. *Publications* **2016**, *4*, 9. [CrossRef]
37. Allen, M.; Poggiali, D.; Whitaker, K.; Marshall, T.R.; Kievit, R. Raincloud plots: A multi-platform tool for robust data visualization. *PeerJ Prepr.* **2018**, *6*, e27137. Available online: <https://doi.org/10.7287/peerj.preprints.27137v1> (accessed on 15 April 2021). [CrossRef]
38. Allen, M.; Poggiali, D.; Whitaker, K.; Marshall, T.R.; Van Langen, J.; Kievit, R.A. Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Res.* **2021**, *4*, 63. [CrossRef]
39. Pagliaro, M. Purposeful Evaluation of Scholarship in the Open Science Era. *Challenges* **2021**, *12*, 6. [CrossRef]
40. Lalli, R. 'Dirty work', but someone has to do it: Howard P. Robertson and the refereeing practices of Physical Review in the 1930s. *Notes Rec. R. Soc. J. Hist. Sci.* **2016**, *70*, 151–174. [CrossRef]
41. Horrobin, D.F. The Philosophical Basis of Peer Review and the Suppression of Innovation. *JAMA* **1990**, *263*, 1438–1441. [CrossRef] [PubMed]
42. Dudley, R.G. The changing landscape of open access publishing: Can open access publishing make the scholarly world more equitable and productive? *J. Libr. Sch. Commun.* **2021**, *9*, 2345. [CrossRef]
43. Demir, S.B. Predatory journals: Who publishes in them and why? *J. Inf.* **2018**, *12*, 1296–1311. [CrossRef]
44. Laakso, M. Green open access policies of scholarly journal publishers: A study of what, when, and where self-archiving is allowed. *Scientometrics* **2014**, *99*, 475–494. [CrossRef]
45. Björk, B.-C.; Laakso, M.; Welling, P.; Paetau, P. Anatomy of green open access. *J. Assoc. Inf. Sci. Technol.* **2014**, *65*, 237–250. [CrossRef]
46. Maddi, A. Measuring open access publications: A novel normalized open access indicator. *Scientometrics* **2020**, *124*, 379–398. [CrossRef]
47. The review of referee 2 for the present work.
48. Serghiou, S.; Ioannidis, J.P.A. Altmetric scores, citations, and publication of studies posted as preprints. *JAMA* **2018**, *319*, 402–404. [CrossRef] [PubMed]
49. MDPI. To Access the Updated Ranking of the World's Top 20 Publishers by Number of Articles and by Number of Journals Published Each Year. 2021. Available online: <https://www.scilit.net/rankings> (accessed on 14 May 2021).
50. Preprint Sharing; Springer: Berlin/Heidelberg, Germany, 2021; Available online: <https://www.springer.com/gp/open-access/preprint-sharing/16718886> (accessed on 14 May 2021).
51. Bar-Ilan, J.; Halevi, G.; Milojević, S. Differences between Altmetric data sources—A case study. *J. Altmetrics* **2019**, *2*, 1. [CrossRef]