



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Contents lists available at ScienceDirect

## European Journal of Internal Medicine

journal homepage: [www.elsevier.com/locate/ejim](http://www.elsevier.com/locate/ejim)

## Letter to the Editor

## 'Paperdemic' during the COVID-19 pandemic

## Dear editor

The 2019 novel coronavirus disease (COVID-19) has spread globally with substantial health and socioeconomic impacts. The scientific community has timely responded to this global crisis by launching calls for COVID-19 research and prioritizing their publications. Journals have accelerated the peer-review process to ensure rapid dissemination of the findings and therefore have received an overwhelming number of papers on the COVID-19 pandemic. A PubMed-based search identified more than 100,000 COVID-19 articles published in 2020, leading to a ~10% increase in the total publication numbers over 2019 [1]. The unprecedented speed and volume of publishing on COVID-19 have greatly contributed to advancing our knowledge of this pandemic and to timely guiding medical management and policy-making. However, the quality and integrity of COVID-19 papers and trails have frequently been questioned, which has led to a large number of corrections and retractions of published materials and created confusion among health-care communities [2,3]; on the other hand, those huge number of COVID-19 papers affected journals to publish non-COVID-19 articles in time and resulted in heavily overloaded work for production teams. This is largely because, as Preiser wrote, some researchers embarked on opportunistic study outside their own specialist fields under this current research ecosystem [4]. The phenomenon of the outbreak of publishing enormous numbers of articles placing a conspicuous burden on journals and yet having little scientific significance is named "paperdemic" (i.e., a pandemic of paper), which will be quantitatively demonstrated in this study from the perspective of scientific impact.

To examine this issue, we collected COVID-19-related papers published in English-language journals from January 1, 2020 to December 31, 2021. A total of 191,368 COVID-19-related papers published in 14,388 academic journals were included, with a fast growth rate after March 2020 (Fig. 1A). It is well known that the scientific impact among publications is imbalanced [5]. Most citation-based measurements for a broad array of scientific disciplines commonly show a highly skewed distribution by fitting the power laws [6]. Studies have also determined the degree of inequality of citations using the Gini coefficient and found disparities in citation distributions [7]. Based upon the fact that an expectably extreme uneven distribution of citations might indicate most of the papers having an inappreciable scientific impact under the dramatic surge of COVID-19 articles, we therefore quantify the degree of inequality in the citation counts of these articles at three different levels, including the paper, journal and discipline, to investigate the phenomenon of paperdemic.

At the individual paper level, we found 80% of citations came from merely 8.3% of all COVID-19 articles. The Gini coefficient value of 0.90 for citation inequality was greater than that for non-COVID-19-related papers (0.84) during the same period and was significantly larger than

that for previous studies [5], which had a Gini coefficient of 0.61 for articles published within one year in other fields. In addition, the normalized citation measurement, calculated by dividing the citation counts that the paper received with the time after its publication, also yielded a similar result, with a Gini coefficient of 0.88 (Fig. 1B), indicating a higher level of inequality in the scientific impact per individual COVID-19 article. At the journal level, the results showed that 80% of the COVID-19 publications came from 21% (3021 out of 14,388) of related journals. Comparatively, 80% of the total scientific impact of COVID-19 articles came from 4% of the journals, which was more unbalanced than the widely mentioned 80/20 rule [8]. The citation distributions were highly dispersed among journals, with Gini coefficients of 0.94, larger than that of the distribution of the number of COVID-19 articles (0.75) (Fig. 1C). The degree of inequality in the distribution of citation was higher than that of other previous studies that found Gini coefficients of 0.58 and 0.65 for the citations earned by journals.

Given that the top 20% of the most cited papers gained 93% of the citations, the inequality of scientific impact was further identified for the most impacted articles by fitting the power laws in the right tail of the citation distributions (Fig. 1D). The power-law exponents were estimated to be 2.4, and was substantially lower than those found in the earlier literatures (2.9–5.3). Previous studies have found that the exponent of power-law fit decreased with time [9], implying an even greater skewness to the right in the future for the scientific impact of COVID-19 articles. The power-law exponent of the citation distribution at the journal level was 2.0, smaller than those for the individual COVID-19 articles, meaning that a small number of journals contributed most of the citations (Fig. 1E). In addition, the top 20% of journals, ranked by the number of published COVID-19 articles, held approximately 92% of the total citations.

The relationship between the proportion of COVID-19-related articles in an individual journal (P1 in Fig. 1F) and the relative amount of total citations among all journals (P2 in Fig. 1D) implied a notable imbalance between the volumes occupied and the corresponding scientific impact among journals. Although few reputable journals yielded a relatively high scientific impact (i.e. larger P2) with a comparable percentage of COVID-19 articles (i.e. smaller P1), for instance, The New England Journal of Medicine, The Journal of the American Medical Association, and The Lancet, a high Gini coefficient (>0.8) was commonly found within these journals.

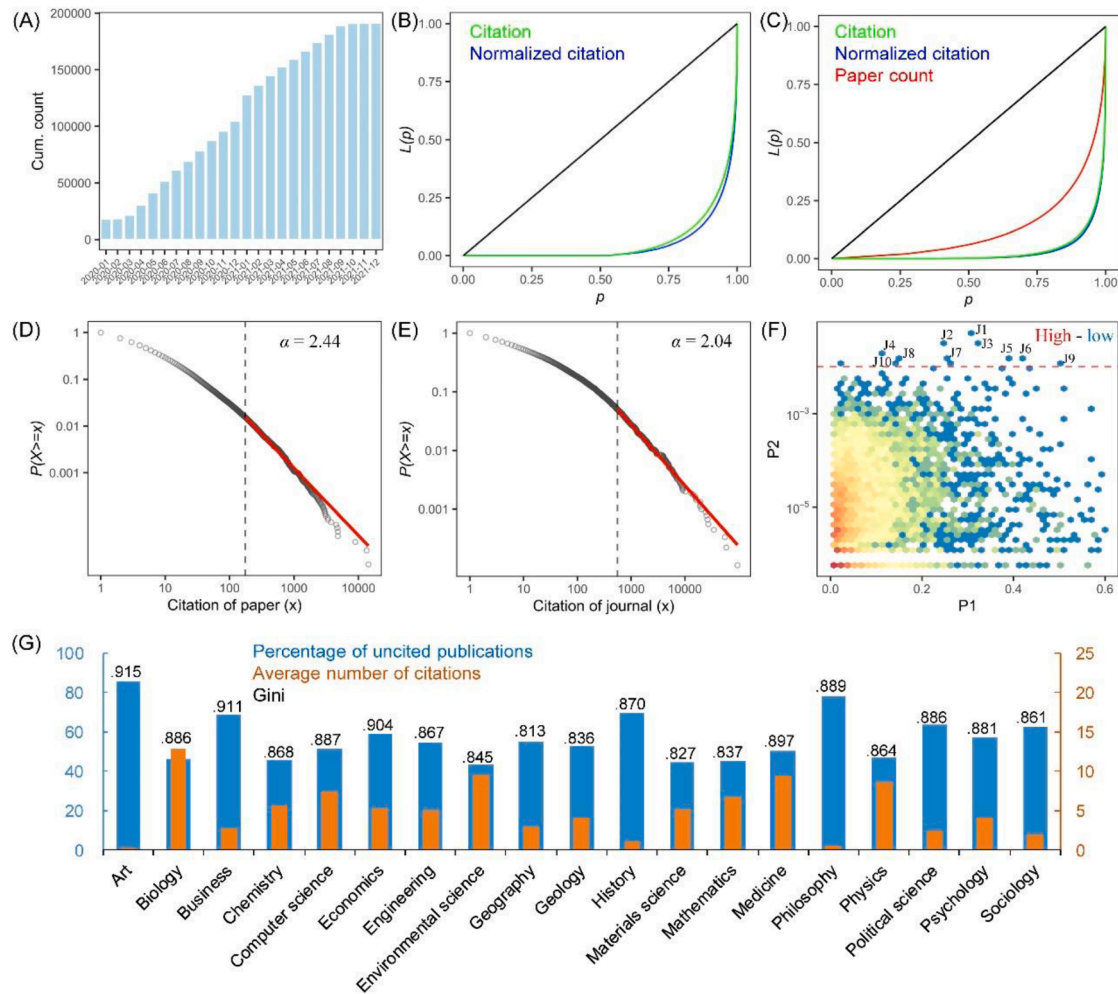
Furthermore, the phenomenon of the paperdemic was likely arising across a range of scientific disciplines, particularly in non-medical and non-biological fields (Fig. 1G). High proportions of uncited articles and low average citation counts occurred in Art, Philosophy, History, Business, Political science, and Sociology. 13 of 19 disciplines showed more than half of published literatures having not been cited up to now. Although the fields of Biology and Medicine exhibited relatively low

<https://doi.org/10.1016/j.ejim.2022.09.030>

Received 1 August 2022; Received in revised form 27 September 2022; Accepted 30 September 2022

Available online 3 October 2022

0953-6205/© 2022 Published by Elsevier B.V. on behalf of European Federation of Internal Medicine.



**Fig. 1.** (A) Cumulative number of COVID-19 papers included in this study for the period from January 1, 2020, to December 31, 2021. (B) and (C) show the Lorenz curves used in Gini coefficients for paper-based and journal-based assessments, respectively. P stands for the cumulative proportion of papers and journals, and  $L(P)$  represents the cumulative proportion of citations, normalized citations, and counts of papers. (D) and (E) exhibit the power-law fitting on the right tail of the citation distributions for the paper-level and journal-level investigations, respectively. The y-axes in (D) and (E) represent the cumulative probability of receiving exactly  $x$  citations and at least  $x$  citations for COVID-19 papers and journals, respectively. (E) shows the relationship between the proportion of COVID-19-related articles in an individual journal (i.e.,  $P_1$ ) and the relative contribution of total citations (i.e.,  $P_2$ ) among all journals. J1- The New England Journal of Medicine; J2- The Journal of the American Medical Association; J3- The Lancet; J4- Science; J5- Lancet Infectious Diseases; J6- The British Medical Journal; J7- Clinical Infectious Diseases; J8- Nature, J9- Journal of Medical Virology, and J10- Radiology. (G) Citation characteristics of COVID-related publications for 19 scientific disciplines (Papers published after October 1, 2021 were excluded for the uncited publication calculation).

percentage of un-cited literatures and high value in average citation number, they had a high Gini coefficient ( $>0.8$ ), which commonly existed in all scientific disciplines.

In summary, our results identified extreme inequalities in the citation index among COVID-19 publications at the paper, journal and discipline levels. This implies that a sizeable COVID-19 literature published in a very short period left unattended along with most of journals and fields, and hence raises major concerns regarding COVID-19 related publications. The need for ensuring rigorous high-quality publications is highlighted for journals particularly during the pandemic. Editors are suggested to ensure the high priority research deserving accelerated publishing process. In contrast, the publication of outsized papers with repeated or similar topics and analyses should be carefully considered to avoid the next wave of the paperdemic, as it is pointed out by a recent astonishing finding [10] that the canonical progress of a scientific field could be slowed by a deluge of papers.

#### Declaration of Competing Interests

None declared.

#### Acknowledgments

This work was supported by the National Natural Science Foundation of China (42071374).

#### References

- [1] Sloane PD, Zimmerman S. The impact of the COVID-19 pandemic on scientific publishing. *J Am Med Dir Assoc* 2021;22:484–8.
- [2] Glasziou PP, Sanders S, Hoffmann T. Waste in covid-19 research. *Brit Med J* 2020; 369:m1847.
- [3] Dal-Re R, Mahillo-Fernandez I. Waste in COVID-19 clinical trials conducted in western Europe. *Eur J Intern Med* 2020;81:991–3.
- [4] Presier W, Presier R. Academic publishing in pandemic times. *S Afr J Sci* 2020;116: 9–10.
- [5] Barabási AL, Song C, Wang D. Handful of papers dominates citation. *Nature* 2012; 491:40.
- [6] Brzezinski M. Power laws in citation distributions: evidence from Scopus. *Scientometrics* 2015;103:213–28.
- [7] Nielsen MW, Andersen JP. Global citation inequality is on the rise. *Proc Natl Acad Sci USA* 2021;7:e2012208118.
- [8] Richard K. *The 80/20 principle: the secret of achieving more with less*. Nicholas Brealey Publishing; 2001.

- [9] Golosovsky M. Power-law citation distributions are not scale-free. *Phys Rev E* 2017;96:032306.
- [10] Chu JSG, Evans JZ. Slowed canonical progress in large fields of science. *Proc Natl Acad Sci USA* 2021;118:e2021636118.

- <sup>b</sup> College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China
- <sup>c</sup> Shanghai Jiao Tong University, Shanghai, China

Yang Yang<sup>a,b</sup>, Na Zhao<sup>a,b,\*</sup>, Ting Ma<sup>a,b</sup>, Ze Yuan<sup>a,b</sup>, Cheng Deng<sup>c</sup>  
<sup>a</sup> State Key Laboratory of Resources and Environment Information System,  
Institute of Geographic Sciences and Natural Resources Research, Chinese  
Academy of Sciences, Beijing, China

\* Corresponding author.  
E-mail address: [zhaon@reis.ac.cn](mailto:zhaon@reis.ac.cn) (N. Zhao).