



RESEARCH ARTICLE

The strain on scientific publishing

Mark A. Hanson¹ , Pablo Gómez Barreiro² , Paolo Crosetto³ , and Dan Brockington^{4,5,6} 

¹Centre for Ecology and Conservation, Faculty of Environment, Science and Economy,
University of Exeter, Penryn Campus, Penryn, UK

²Royal Botanic Gardens, Kew, Richmond, UK

³Université Grenoble Alpes, INRAE, CNRS, Grenoble, France

⁴Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autònoma de Barcelona, Barcelona, Spain

⁵Institución Catalana de Investigación y Estudios Avanzados (ICREA), Barcelona, Spain

⁶El Departament de Dret Privat, Universitat Autònoma de Barcelona, Barcelona, Spain

an open access  journal



Citation: Hanson, M. A., Barreiro, P. G., Crosetto, P., & Brockington, D. (2024). The strain on scientific publishing. *Quantitative Science Studies*, 5(4), 823–843. https://doi.org/10.1162/qss_a_00327

DOI:
https://doi.org/10.1162/qss_a_00327

Peer Review:
https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00327

Supporting Information:
https://doi.org/10.1162/qss_a_00327

Received: 11 November 2023
Accepted: 28 July 2024

Corresponding Author:
Mark A. Hanson
m.hanson@exeter.ac.uk

Handling Editor:
Vincent Larivière

Copyright: © 2024 Mark A. Hanson, Pablo Gómez Barreiro, Paolo Crosetto, and Dan Brockington. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.



Keywords: article processing charge, impact factor, open access, Scopus, Web of Science

ABSTRACT

Scientists are increasingly overwhelmed by the volume of articles being published. The total number of articles indexed in Scopus and Web of Science has grown exponentially in recent years; in 2022 the article total was ~47% higher than in 2016, which has outpaced the limited growth—if any—in the number of practicing scientists. Thus, publication workload per scientist has increased dramatically. We define this problem as “the strain on scientific publishing.” To analyze this strain, we present five data-driven metrics showing publisher growth, processing times, and citation behaviors. We draw these data from web scrapes, and from publishers through their websites or upon request. Specific groups have disproportionately grown in their articles published per year, contributing to this strain. Some publishers enabled this growth by hosting “special issues” with reduced turnaround times. Given pressures on researchers to “publish or perish” to compete for funding, this strain was likely amplified by these offers to publish more articles. We also observed widespread year-over-year inflation of journal impact factors coinciding with this strain, which risks confusing quality signals. Such exponential growth cannot be sustained. The metrics we define here should enable this evolving conversation to reach actionable solutions to address the strain on scientific publishing.

1. INTRODUCTION

Academic publishing has a problem. The last few years have seen an exponential growth in the number of peer-reviewed journal articles, which has not been matched by the training of new researchers who can vet those articles. Editors are reporting difficulties in recruiting qualified peer reviewers (Fox, Albert, & Vines, 2017; Peterson, Orticio, & Nugent, 2022), and scientists are overwhelmed by the volume of new articles being published (Parolo, Pan et al., 2015; Severin & Chataway, 2021). We will call this problem “the strain on scientific publishing.”

Part of this growth may come from inclusivity initiatives or investment in the Global South, which make publishing accessible to more researchers (Maher, Aseffa et al., 2020; Nakamura, Soares et al., 2023). Parallel efforts have also appeared in recent years to combat systemic biases in scientific publishing (Langin, 2023; Liu, Brown, & Sabat, 2019; Meijaard, Cardillo et al., 2015), including positive-result bias (Mlinarić, Horvat, & Šupak Smolčić, 2017). To the extent that the growth comes from such initiatives, it is welcome and should be accommodated.

However, growth will become strain if it compromises the ability of scientists to be rigorous when vetting information (Hofseth, 2018). If scientific rigor is allowed to slip, it devalues the term “science” (Sarewitz, 2016). Recent controversies already demonstrate this threat, as research paper mills operating within publishing groups have caused mass article retractions (Abalkina, 2023; Bishop, 2023; Candal-Pedreira, Ross et al., 2022; Else & Van Noorden, 2021), alongside renewed calls to address so-called “predatory publishing” (Grudniewicz, Moher et al., 2019).

To understand the forces that contribute to this strain, we first present a simple schematic to describe scientific publishing. We then specifically analyze publishers, as their infrastructures regulate the rate at which growth in published articles can occur. To do this, we identify five key metrics that help us to understand the constitution and origins of this strain: growth in total articles and special issues, differences in article turnaround times or rejection rates, and a new metric based on citation behavior that we call “impact inflation.”

These metrics should be viewed in light of publisher business models. First, there is the more classic subscription-based model generating revenue from readers. Second, there is the “gold open access” model, which generates revenue through article processing charges that authors pay. In both cases publishers can act either as for-profit or not-for-profit organizations. We therefore consider whether aspects of either of these business models are contributing to the strain.

By performing a comparative analysis combining multiple metrics, we find that strain is not strictly tied to any one publisher business model, although some behaviors are associated with specific gold open access publishers. We argue that existing efforts to address this strain are insufficient. We highlight specific areas needing transparency, and actions that publishers, researchers, and funders can take to respond to this strain. Our study provides essential data to inform the existing conversation on academic publishing practices.

2. THE LOVE TRIANGLE OF SCIENTIFIC PUBLISHING: A CONCEPTUAL FRAMEWORK

The strain on scientific publishing is the result of interactions between three sets of players: publishers, funders, and researchers.

Publishers want to publish as many papers as possible, subject to a quality constraint, which varies from brand to brand. This is because publishers advertise total articles published as a desirable quality that augments their brand (File S1), as bargaining chips to negotiate subscription fees (Khelifaoui & Gingras, 2022; Shu, Mongeon et al., 2018), and, when gold open access is involved, publishers earn revenue in direct proportion to total articles published (Butler, Matthias et al., 2023); all items labeled with the prefix “S” may be found in the Supplementary material. Journal brands (e.g., “Nature”) can also be used to market other services, or even other journals (Khelifaoui & Gingras, 2022). Publications themselves give researchers a “stamp of quality” that researchers use to advance their own goals. The quality of a stamp is often determined by journal-level prestige metrics, such as the Clarivate journal Impact Factor (IF), or Scopus Scimago Journal Rank (SJR) (Garfield, 2006; Guerrero-Bote & Moya-Anegón, 2012), and ultimately by association with the quality of published papers. Publishers compete among each other to attract the most and/or the best papers.

Funders (e.g., universities, funding agencies) use “stamps” from the science publication market as measures of quality to guide their decisions on whom to hire and fund. Ultimately, money from funders supports the whole market, and funders want cost-effective, stable, and informative signals to help guide their decisions. Some funders are experimenting with ways to reward researchers independent from publisher stamps (e.g., narrative CVs (DORA, 2022)),

though others use journal stamps as signals to determine researcher promotion or salary (e.g., Quan, Chen, & Shu, 2017).

Finally, due to “publish or perish” pressures, researchers are incentivized to publish as many papers in prestigious journals as possible (Quan et al., 2017; Sarewitz, 2016). If publications are among the main outputs used by funders to gauge researcher productivity then researchers must be productive, which becomes synonymous with “publish,” to secure employment, promotion, and funding. Researchers are also the backbone of science, acting as authors that generate articles, but also as referees and editors during peer review for publishers and funders, almost always for free. As a result, they help influence the administering of publisher stamps of quality. More altruistically, they help ensure the quality of science in their field.

The incentives for publishers and researchers to increase their output drive growth. This is not problematic *per se*, but there will be a trade-off between the volume of work being produced and its quality. The difficulty is that “quality” is hard to define (Garfield, 2006; Guerrero-Bote & Moya-Anegón, 2012; Thelwall, Kousha et al., 2023), and some metrics are at risk of abuse per Goodhart’s law: “when a measure becomes a target, it ceases to be a good measure” (Fire & Guestrin, 2019). For instance, having many citations may indicate that an author, article, or journal, is having an impact. But citations can be gamed through self-citing or coordinated “citation cartels” (Abalkina, 2023; Bishop, 2023; Fong, Patnayakuni, & Wilhite, 2023).

Collectively, the push and pull by the motivations of these players defines the sum product of the scientific publishing industry.

3. MATERIALS AND METHODS

3.1. Description of Publisher Data

We produced five metrics of publisher practice that describe the total volume of material being published, or that affect the quality of publisher “stamps.” We focused our analyses on the last decade of publication growth, with special attention paid to the period 2016–2022, as pre-2016 some data types were less available. We used the Scopus database (via Scimago (Scimago, 2023)) filtered for journals indexed in both Scopus and Web of Science. Our focus on journals indexed by both Scopus and Web of Science describes a more conservative measure of journal data compared to other possible data sets (e.g., Dimensions, OpenAlex), but benefits from focusing only on journals that pass stricter indexing criteria. We further assembled journal/article data by scraping information in the public domain from web pages, and/or following direct requests to publishers (Table 1). These metrics are:

- Total articles indexed in both Scopus and Web of Science
- Share of articles appearing in special issues
- Article turnaround times from submission to acceptance
- Journal rejection rates as defined by publishers
- A new metric we call “impact inflation,” informed by journal citation behaviors

3.2. Data Collection

3.2.1. Publisher and journal-level data: Publisher selection

Due to technical limits of web-scraping, we focused our analyses of special issue proportions, turnaround times, and rejection rates, on only a subset of publishers and articles. We chose publishers according to their size and importance, and to get a sufficient variety of business

Table 1. Summary of web scraped data informing share of special issue articles and turnaround times. For some publishers, the number of web scraped journals or articles with turnaround time data exceeds the totals from our Scimago data set (noted with *). This is because, in the web-scraped data set, we included all journals by a given publisher, even if they were not indexed, or indexed by only one of Scopus or Web of Science.

Publisher	Web scraped journals included	Articles with turnaround times ¹	Total journals (Scimago)	Total articles (Scimago) ¹
BMC*	289	344,501	213	241,493
Elsevier	376	531,580	1579	2,988,422
Frontiers	44	291,017	49	329,370
Hindawi*	220	226,612	161	155,396
MDPI	98	838,448	152	840,518
Nature*	144	360,855	111	346,845
PLOS	12	243,398	7	148,404
Springer	1,259	1,371,405	1,589	1,378,386
Taylor & Francis	1,063	512,438	1,160	525,029
Wiley	1,257	890,174	1,467	1,450,487

¹ Time period 2016–2022.

models. We included Elsevier, Wiley-Blackwell (Wiley), Springer, Nature, and Taylor & Francis, which are among the largest long-established academic publishers, and which use a mixture of subscription and gold open access licenses. Multidisciplinary Publishing Institute (MDPI), Frontiers Media (Frontiers), and Hindawi are younger for-profit gold open access publishers well known for their business model that publishes many articles through special issues. In contrast, BMC is a for-profit gold open access publisher that operates hundreds of journals but does not publish many special issues by comparison. Finally, PLOS is a not-for-profit gold open access publisher, and among the largest publishers in terms of articles per journal per year.

Previous analyses have grouped journals under the umbrella of their parent company ownership (Butler et al., 2023; Larivière, Haustein, & Mongeon, 2015). Here we have aggregated publisher labels according to their corporate branding in Scimago: For instance, Elsevier BV, Elsevier Ltd, and similar were aggregated as “Elsevier,” or Springer GmbH & Co and Springer International Publishing AG as “Springer.” However, we have not grouped subsidiary brands under the umbrella of their parent company (e.g., Nature and BMC journals are owned by the conglomerate “Springer-Nature”). We did this because the behavior of these subsidiary groups is different from the parent company (sometimes radically so). Certain subsidiary publisher brands further define meaningful groups for comparison (e.g., BMC, Hindawi). The status of publisher brands has also changed over the years: a majority share in Frontiers was once owned by Nature Publishing Group (Enserink, 2015), and Hindawi was purchased by Wiley in 2021, who in 2023 decided to “sunset” the brand following paper mill scandals (Bishop, 2023). If we merged these brands with their parent companies we would obscure and confuse relevant data comparisons. For instance, Hindawi would merge into Wiley in 2022 causing dropout of Hindawi from the 2022 data set and a spike in Wiley’s total articles and use of special issue publishing by virtue of a merger, and not through a change in behavior of pre-existing Wiley journals.

3.2.2. Publisher and journal-level data: Data curation

Data for total articles published per year come from a subset of the Scopus database obtained from Scimago (2023). Historical data (1999–2022) for total number of articles, total citations per document over 2 years, the SJR metric, and total references per document were obtained from the Scimago web portal (<https://www.scimagojr.com/journalrank.php>). Scimago yearly data were downloaded with the “only WoS journals” filter applied to ensure the journals we include here were indexed by both Scopus and Web of Science (Clarivate). Within-journal self-citation rate was obtained from Scimago via web scraping.

Historical Impact Factor data (2012–2022) for a range of publishers (16,174 journals across BMC, Cambridge University Press, Elsevier, Emerald Publishing Ltd., Frontiers, Hindawi, Lippincott, MDPI, Springer, Nature, Oxford University Press, PLOS, Sage, Taylor & Francis, and Wiley-Blackwell) were downloaded from Clarivate. Due to the download limit of 600 journals per publisher, these IFs represent only a subset of all IFs available.

Rejection rates were collected from publishers in a variety of ways:

- obtained from online available publisher reports (Frontiers: <https://progressreport.frontiersin.org/peer-review/>);
- given by publishers upon request (PLOS, Taylor & Francis); and
- web scraping of publicly available data extracted from the journal or company websites (MDPI, Hindawi, Elsevier via <https://web.archive.org/web/20220401133911/https://journalinsights.elsevier.com/journals/1072-7515>).

Frontiers rejection rate data lack journal-level resolution, and are instead the aggregate of the whole publisher corpus per year.

Data on publisher values, goals, motivations, and aspirations (see File S1) were obtained from annual reports, press releases, and reports to shareholders.

3.2.3. Article-level data

Several methods were used to obtain article submission and acceptance times, which were used to calculate turnaround times, here defined as the time taken from submission to formal acceptance. Turnaround times advertised by publishers often refer to the time taken until first decision (either accept or reject), which is affected by the rate of submissions and the rate of rejection, two metrics that are not readily observed. By focusing on turnaround times from submission to acceptance, we strictly assess the time spent considering papers deemed worth publishing, which better reflects the rigor of the peer review process for a given publisher, is a comparable metric across publishers, and is readily available on articles' websites. PLOS, Hindawi, and Wiley's turnaround times were extracted directly from their corpus. The latter was shared with the authors by Wiley upon request, while the corporuses of PLOS (<https://plos.org/text-and-data-mining/>) and Hindawi (<https://www.hindawi.com/hindawi-xml-corpus/>) are available online. For other turnaround times, we used web scraping, extracting the publicly reported date of submission and date of acceptance from article web pages.

Annotation of whether articles were contained in “special issues” required some degree of manual data curation. Broadly, if publishers defined articles as belonging to a “special issue” or other article collection heading (e.g., “Theme Issues,” “Collections,” or “Topics”), we considered those articles to be special issue articles. Per definitions from the Committee on Publication Ethics (COPE, 2023), these articles are often handled by guest editors, but special

issues can also be handled by in-house editors. Special issues also vary in format and editorial practice across publishers. As such, we have used an inclusive approach to describing article collections as special issues, ensuring we do not filter out article collections for vague reasons.

3.2.4. Web scraping

Web scraping scripts considered a publisher's website formatting and article metadata formatting, and so required manual curation to tailor them to each publisher. BMC, Frontiers, MDPI, Nature, and Springer data were obtained via web scraping of individual articles and collecting data in "article information"-type sections. Taylor & Francis turnaround times were obtained via CrossRef (CrossRef, 2023) by filtering all available ISSNs from Scimago. To obtain Elsevier turnaround times we first extracted all Elsevier related ISSNs from Scimago, queried these in CrossRef to obtain a list of DOIs, and then web scraped the data from those articles. We also collected information on whether Elsevier articles were part of special issues during our web scraping. However, the resulting data were unusually spotty and incomplete: For instance, we had journals with a total of only one article with data on special issue status, which would falsely suggest that 100% of articles in that journal were special issue articles. Ultimately, we did not include Elsevier in our analysis of special issue articles.

While web scraping had to be tailored to each publisher website, the core strategy was consistent across web-scraped publishers. In summary, URLs of published articles were obtained in order to be able to download their HTML or XML code and inspect nodes containing information relevant to editorial times and special issues. Extracted data was then formatted to be consistent and allow publisher-to-publisher comparisons.

3.2.5. Global researcher statistics

Total PhD graduate numbers were obtained from the Organisation for Economic Co-operation and Development (OECD; 37 countries at the time of data collection: <https://stats.oecd.org>, data up to 2020). Other sources were consulted to complement OECD data with PhD graduate data for China and India (National Science Foundation, 2022; Zwetsloot, Corrigan et al., 2021) but data for these countries only existed up to 2019 and were not collected using the same parameters as the OECD data. We therefore present PhD data on OECD countries in the main text, and the confirmed and estimated data trends beyond OECD data in the Supplementary material.

We further complemented PhD graduate data with measures of researchers-per-million (full-time equivalent) from the February 2023 release of the UNESCO Science, Technology, and Innovation data set (<https://data.uis.unesco.org>, "9.5.2 Researchers per million inhabitants") and projected these data to 2022 using a linear regression model.

3.2.6. Rejection rates

We defined rejection rate as a function of accepted, rejected, and total submissions, depending on the data that were available for each publisher. Of note, rejection rates are affected by a number of cryptic factors that make direct comparisons complicated. For example, journals may differ in their culture for how long revisions are permitted before a resubmission is classified as a new article.

For Frontiers, we use the formula $1 - (\text{accepted articles} / \text{total submissions})$ to be consistent with data available for other publishers. We defined total submissions for MDPI as the sum of all accepted and rejected articles and calculated rejection rates accordingly. Hindawi report

their rejection rates publicly on journal pages as “acceptance rate,” although the underlying calculation method is not given.

While we could not standardize the methodology used to calculate rejection rates across publishers, we make the assumption that publishers have at least maintained a consistent methodology internally across years. For this reason, while comparing raw rejection rates comes with many caveats, comparing the direction of change itself in rejection rates within groups should be relatively robust to allow comparisons of trends between groups.

3.2.7. Impact inflation

Clarivate’s IF is calculated as the mean total citations per article in articles published within the last 2 years. The SJR is more complex, using a citation network approach that places a higher value on citations between journals of the same general field and limits the contribution to the SJR made by journals through self-citation or citation cartel behaviors. Full details of the SJR metric are given in Guerrero-Bote and Moya-Anegón (2012). Key differences between the two metrics are summarized in Table 2. Importantly, SJR limits the reward in prestige/rank that journals receive from individual sources. Thus, while high rates of self-citation or circular citation are rewarded indefinitely by IF, their contribution to a journal’s SJR is limited. The ratio of IF/SJR can therefore reveal journals whose total citations come from disproportionately few citing journals. We call this ratio “impact inflation,” which applies to both IF/SJR, or the equivalent calculation using Scimago’s IF equivalent “citations per document (Cites-per-doc/SJR).”

We describe publisher characteristics using journal-level values (within journal self-citation rate, IF, Cites-per-doc). Notably, the age of journals is tied to article output, as newer journals publish fewer articles, but can grow to publish thousands of articles annually in later years. New journals are also likely to have lower self-citation rates as, at their inception, they have few articles available to self-cite. Small journals may also have different self-citation behaviors from larger journals if, for example, they cite their own previous papers for things like geographic reasons (e.g., the *Canadian Journal of Fisheries and Aquatic Sciences*). Therefore, we analyze journal data with journal size considered, and filtered for established journals whose articles receive a total of at least 1,000 annual citations for self-citation analyses. This was especially important for comparisons at the publisher level, as some publishers have increased their number of journals substantially in recent years (Figure S1), meaning a large fraction of their journals are relatively young and less characteristic of the publisher’s trends according to their better-established journals.

Table 2. Methodological differences between SJR and impact factor. Table content adapted from Guerrero-Bote and Moya-Anegón (2012).

	SJR	Impact factor
Database	Scopus	Web of Science
Citation time frame	3 years	2 years
Self-citation contribution	Limited	Unlimited
Field-weighted citation	Weighted	Unweighted
Size normalization	Citable document rate	Citable documents
Citation networks considered?	Yes	No

High-resolution versions of all the figures can be found at: https://figshare.com/articles/figure/The_strain_on_scientific_publishing_figures_/24203790.

4. RESULTS

4.1. A Few Publishers Disproportionately Contribute to Total Article Growth

There were ~897,000 more indexed articles per year in 2022 (~2.82 million articles) compared to 2016 (~1.92 million articles) (Figure 1A), a year-on-year exponential growth of ~5.6% over this period. To understand the source of this substantial growth, we first divided article output across publishers per Scopus publisher labels (Figure 1B). The five largest publishers by total article output include Elsevier, MDPI, Wiley, Springer, and Frontiers, respectively. However, in terms of strain added since 2016, their rank order changes: Journals from MDPI (~27%), Elsevier (~16%), Frontiers (~11%), Springer (~9.5%), and Wiley (~7.0%) have contributed more than 70% of the increase in articles per year. Elsevier and Springer own a huge proportion of total journals, a number that has also increased over the past decade (Figure S1). As such, we normalized article output per journal to decouple the immensity of groups like Elsevier and Springer from the growth of articles itself. While Elsevier has increased article outputs per journal slightly, other groups, such as MDPI and Frontiers, have become disproportionately high producers of published articles per journal (Figure 1C).

Taken together, groups like Elsevier and Springer have quantitatively increased total article output by distributing articles across an increasing number of journals. Meanwhile groups like MDPI and Frontiers have been exponentially increasing the number of publications handled by a much smaller pool of journals. These publishers reflect two different mechanisms that have promoted the exponential increase in total articles published over the last few years.

4.2. Growth in Articles Published Through “Special Issues”

“Special issues” are distinct from standard articles because they are invited by journals or editors, rather than submitted independently by authors. They also commonly delegate responsibilities to guest editors, whereas editors for regular issue articles are members of their journal’s editorial board, expected to have more experience with their journal’s specific expectations. In recent years, certain publishers have adopted this business model as a route to publish the majority of their articles (Figure 2). This behavior encourages researchers to generate articles specifically for special issues, raising concerns that publishers could abuse this model for profit (Butler et al., 2023; Ioannidis, Pezzullo, & Boccia, 2023; Larivière et al., 2015). Here we describe this growth in special issues for eight publishers for which we could collect data.

Between 2016 and 2022, the proportion of special issue articles grew drastically for Hindawi, Frontiers, and MDPI (Figures S4 and S5). These publishers depend on article processing charges for their revenues, which are paid by authors to secure gold open access licenses. But this special issue growth is not a necessary feature of open access publishing, as similar changes were not seen in other gold open access publishers (i.e., BMC, PLOS). Publishers using both subscription and open access approaches (Nature, Springer, Wiley) also tended to publish small proportions of articles through special issues.

These data show that the strain generated by special issues is not a direct consequence of the rise of open access publishing *per se*, or associated article processing charges. Instead, the dominance of special issues in a publisher’s business model is publisher dependent.

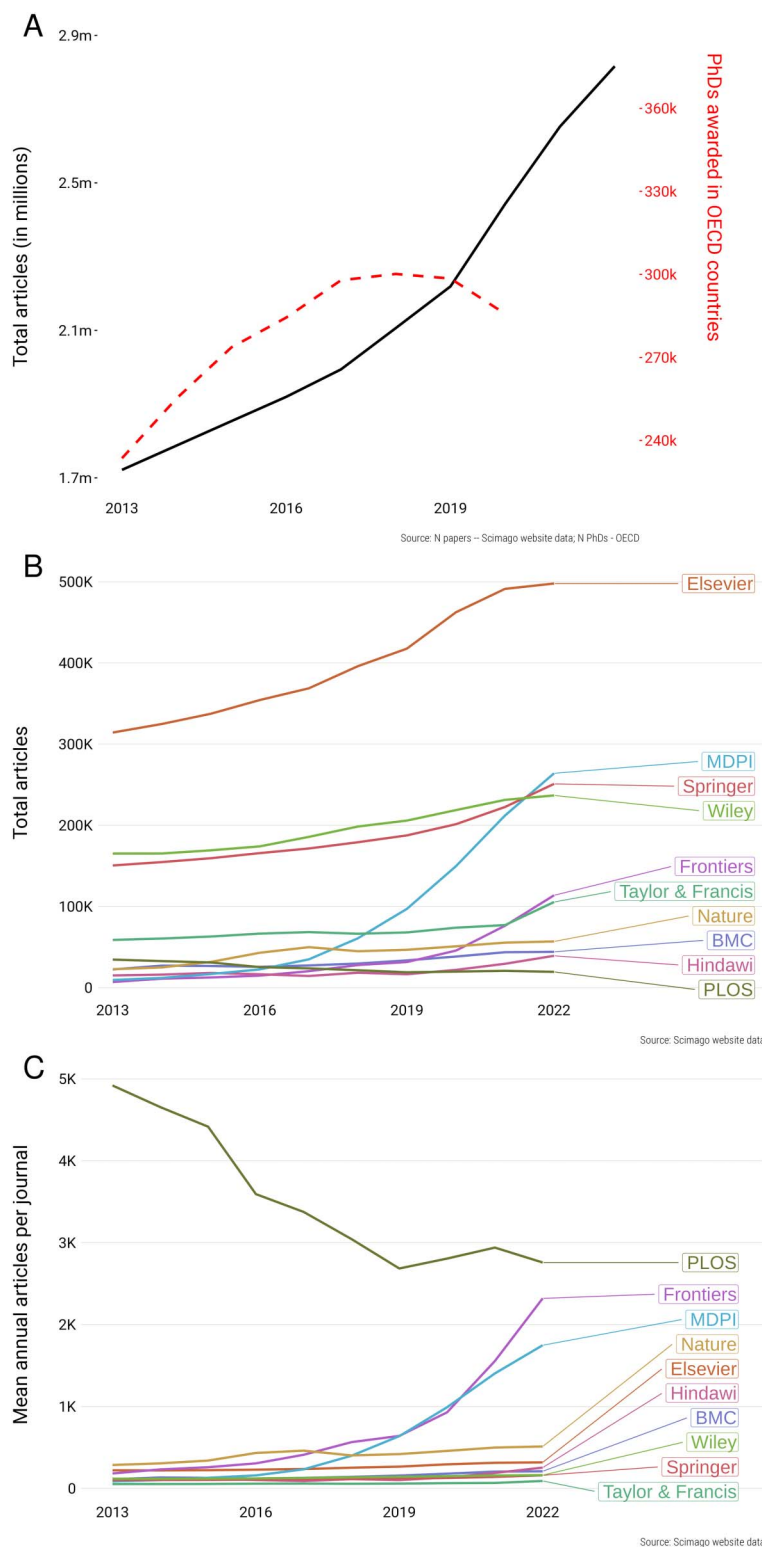


Figure 1. Total article output is increasing. A: Total articles being published per year has increased exponentially, while the number of PhDs being awarded has not kept up. This remains true with the addition of non-OECD countries, or when using global total employed researcher-hours instead of PhD graduates as a proxy for active researchers (Figure S2). B and C: Total articles per year by publisher (B), or per journal per year by publisher (C). Also see growth in journals per publisher (Figure S1) and by size class (Figure S3).

Number of papers published in regular vs special issues, 2016-22

One square = 800 articles



Note: Special issues are called Collections at PLOS and Topics at Frontiers. For MDPI Collections, Sections and Topics not shown. Source: data scraped from the publisher's website

Figure 2. Rise of the special issue model of publishing. Normal articles (blue) and special issue articles (red) over time. Frontiers, Hindawi, and especially MDPI publish a majority of their articles through special issues, including an increase in recent years alongside growth seen in Figure 1 (detailed further in Figures S4 and S5). These data reflect only a fraction of total articles shown in Figure 1, limited due to sampling methodology (Table 1). Elsevier are excluded because of problems of data availability.

4.3. Decreasing Mean, Increasing Homogeneity of Turnaround Times

The time spent in processing an article is colloquially known as its “turnaround time.” Article turnaround times rely on the rapidity of the publisher to provide an editorial decision, and the level of care that researchers take in making edits during revisions before resubmitting the article for consideration. Here we analyze article turnaround times, defined as the time taken from first submission to editorial acceptance. We use this time frame because the time required for revisions can be weeks to months, depending on the field of research and the magnitude or type of revisions required; for instance, minor text changes could be made quickly, but essential experiments requested by reviewers could take months to complete. As a result, turnaround times within journals should be highly heterogeneous if each article is considered and addressed according to its unique needs, as expected of rigorous peer review. Moreover, the average turnaround time is expected to vary from journal to journal according to research field. This is because revision requirements vary by research field: For example, in biology, additional reviewer-requested experiments might take just a few weeks to complete using a fruit fly study system, while experiments with mice might take months.

We analyzed turnaround times between 2016 and 2022 for publications where data were available. We found that average turnaround times vary markedly across publishers. Like others (MDPI, 2021; Oviedo-García, 2021), we found that MDPI had an average turnaround time of ~37 days from first submission to acceptance in 2022, a level they have held since ~2018. This turnaround time is far lower than comparable publishers like Frontiers (72 days) and Hindawi (83 days), which also saw a decline in mean turnaround time between 2020 and 2022. On the other hand, other publishers in our data set had turnaround times of more than 130 days, and, if anything, their turnaround times increased slightly between 2016 and 2022 (Figure 3A).

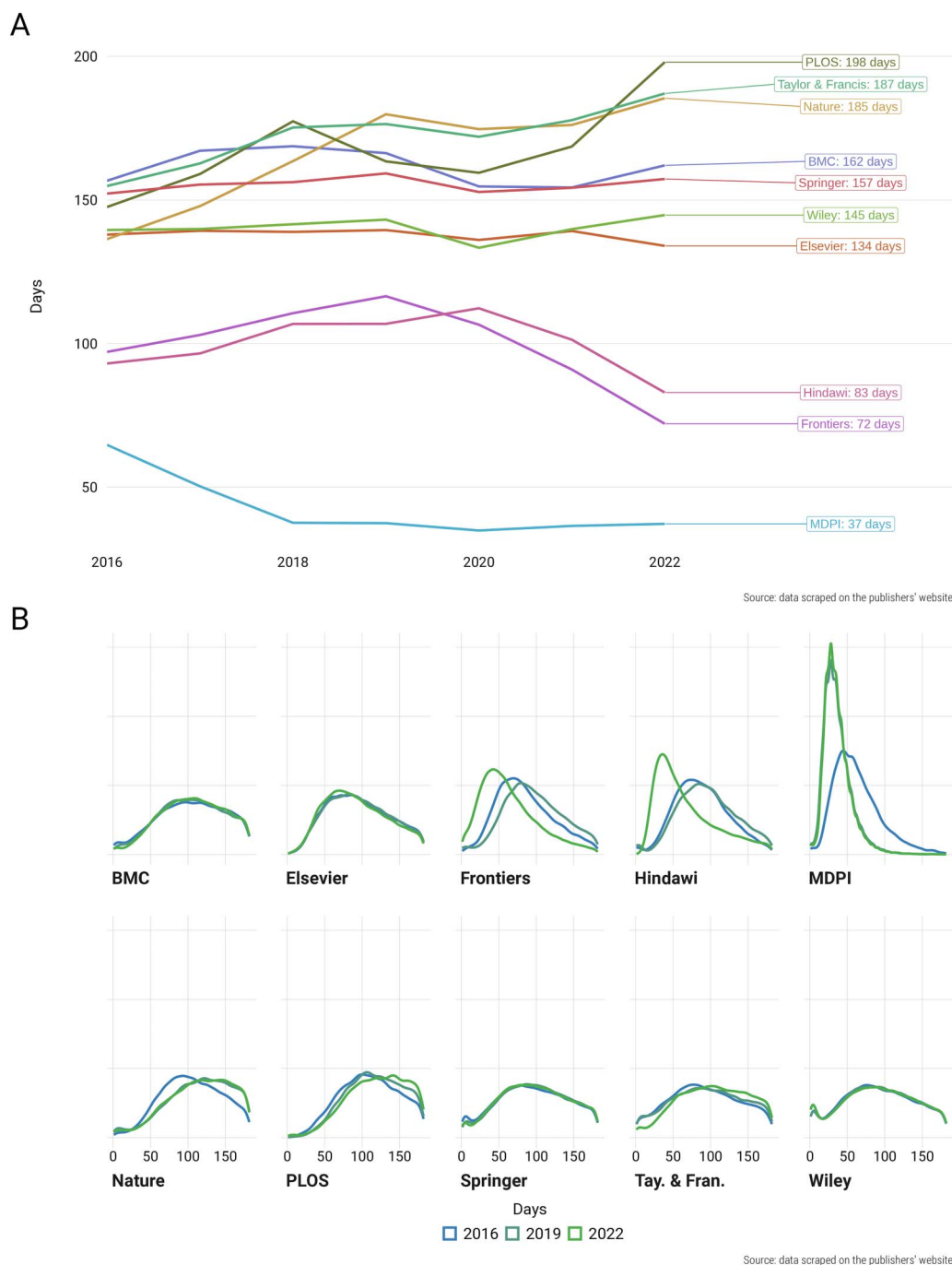


Figure 3. Article turnaround times. A: Evolution of mean turnaround times by publisher. Only articles with turnaround times between 1 days and 2 years were included. This filter was applied to remove data anomalies such as immediate acceptance or missing values that default to Jan 1, 1970 (the “Unix epoch”). B: Article turnaround time distribution curves from 2016–2022, focused on the first six months to better show trends. While most publishers have a right-skewed curve, the three publishers highlighted previously for increased special issue use have a left-skewed curve that only became more extreme over time. These data reflect only a fraction of the total articles shown in Figure 1, limited due to sampling methodology (see Table 1). (Tay. & Fran. = Taylor & Francis. Full descriptive statistics in Table S1.)

The publishers decreasing their turnaround times also show declining variances. Turnaround times for Hindawi, Frontiers, and especially MDPI are becoming increasingly homogeneous (Figure 3B and Figure S6): Articles, regardless of initial quality or field of research, and despite the expectation of heterogeneity, are all accepted in an increasingly similar time frame.

The decrease in mean turnaround times (Figure 3A) also aligns with inflection points for the exponential growth of articles published as part of special issues in Hindawi (2020), Frontiers (2019), and MDPI (2016) (see Figure S4). We therefore asked whether special issue articles are processed more rapidly than normal articles in general. For most publishers, this was indeed the case, even independent of proportions of normal and special issue articles (Figure S7). In the case of Hindawi, a previous study also noted that special issue articles were more likely to involve paper mill activity, suggesting both lower article quality and lower editorial rigor (Bishop, 2023).

In summary, turnaround times differ by publisher, associated with use of the special issue publishing model. Variance in turnaround times also decreases for publishers alongside adoption of the special issue model. These results suggest that special issue articles are typically accepted more rapidly and in more homogeneous time frames than normal articles, which, to our knowledge, has never been formally described.

4.4. Journal Rejection Rates and Trends Are Publisher Specific

If a publisher lowers its article rejection rates, all else being equal, this will lead to more articles being published. Such changes to rejection rate might also mean more lower-quality articles are being published. Peer review is the principal method of quality control that defines science (Grainger, 2007), and so publishing more articles of lower quality may add to strain and detract from the meaning and authority of the scientific process. The relationship between rejection and quality is complex: High or low rejection rates may stem from editorial oversight over scope, or perceived impact. Publishers also define what it means to be “rejected,” creating caveats to comparing raw numbers across publishers.

Rejection rate data are rarely made public, and only a minority of publishers provide these data routinely or shared rejection rates upon request. Using the rejection rate data we could collect, we estimated rejection rates per publisher and asked if they

- change with growth in articles;
- correlate with journal size;
- correlate with journal impact;
- depend on the publisher; or
- predict a journal’s proportion of special issue articles.

We found no clear trend between the evolution of rejection rates and publisher growth (Figure 4A). Focusing on younger journals (≤ 10 years, ensuring fair comparisons) we found no relationship between journal size and reported or calculated 2022 rejection rates (Figure 4B). Finally, citations per document (similar to Clarivate IF) did not correlate with rejection rates (Figure S8), indicating that citations are not a strong predictor. Ultimately, the factor that best predicted rejection rates was the publisher itself: Although both Frontiers and MDPI have similar growth in special issue articles (Figure 2), they show opposite trends in rejection rates over time, and MDPI uniquely showed decreasing rates compared to other publishers (Figure 4A). Raw rejection rates for MDPI in 2022 were also lower than for other publishers. Moreover,

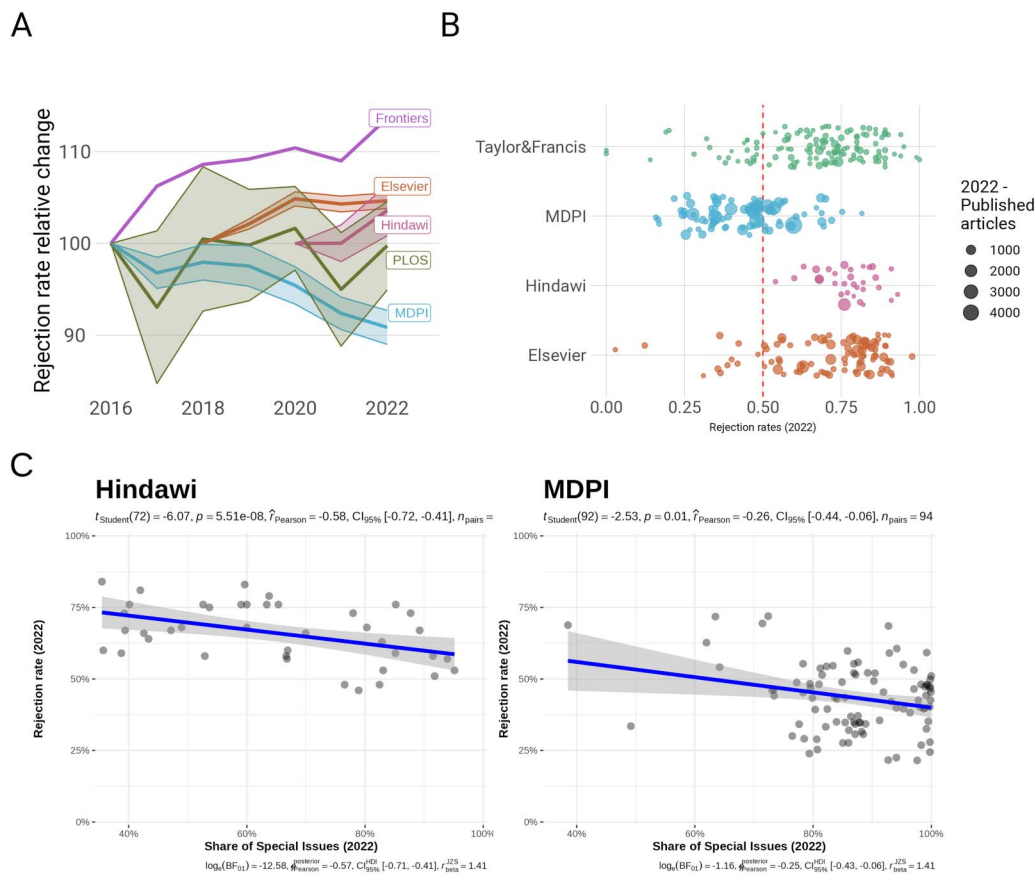


Figure 4. Rejection rates are defined most specifically by publisher. A: Rejection rates increased, decreased, or experienced little change depending on publisher. We estimated publisher rejection rates from varying available data, so we normalized these data by setting the first year on record as “100.” Frontiers data are the aggregate of all Frontiers journals, preventing the plotting of 95% confidence intervals. B: 2022 rejection rates among young journals (less than 10 years old) differ by publisher, but not journal size. C: For two publishers that we could analyze, there was a significant correlation between 2022 journal rejection rates decreasing and the share of articles published through special issues increasing.

Hindawi and MDPI journals with more special issue articles also had lower rejection rates ($p = 5.5 \times 10^{-8}$ and $p = .01$ respectively; data from 2022, Figure 4C and Figure S9), which we could not assess for other publishers.

In summary, we found no general associations across publishers between rejection rate and most other metrics we investigated. Over time or among journals of similar age, rejection rate patterns were largely publisher specific. We did, however, recover a trend that within the limited subset of publishers for which we have data, rejection rates decline with increased use of special issue publishing.

4.5. Disproportionately Inflated Impact Factor Affects Select Publishers

Among the most important metrics of researcher impact and publisher reputation are citations. For journals, the Clarivate two-year IF reflects the mean citations per article in the 2 preceding years. Here we found that IF has increased across publishers in recent years (Figures S10 and S11). Explaining part of this IF inflation, we observed an exponential increase in total references per document between 2018 and 2021 (Figure S12, and see Neff & Olden, 2010). However, we previously noted that IF is used as a “stamp of quality” by both researchers and

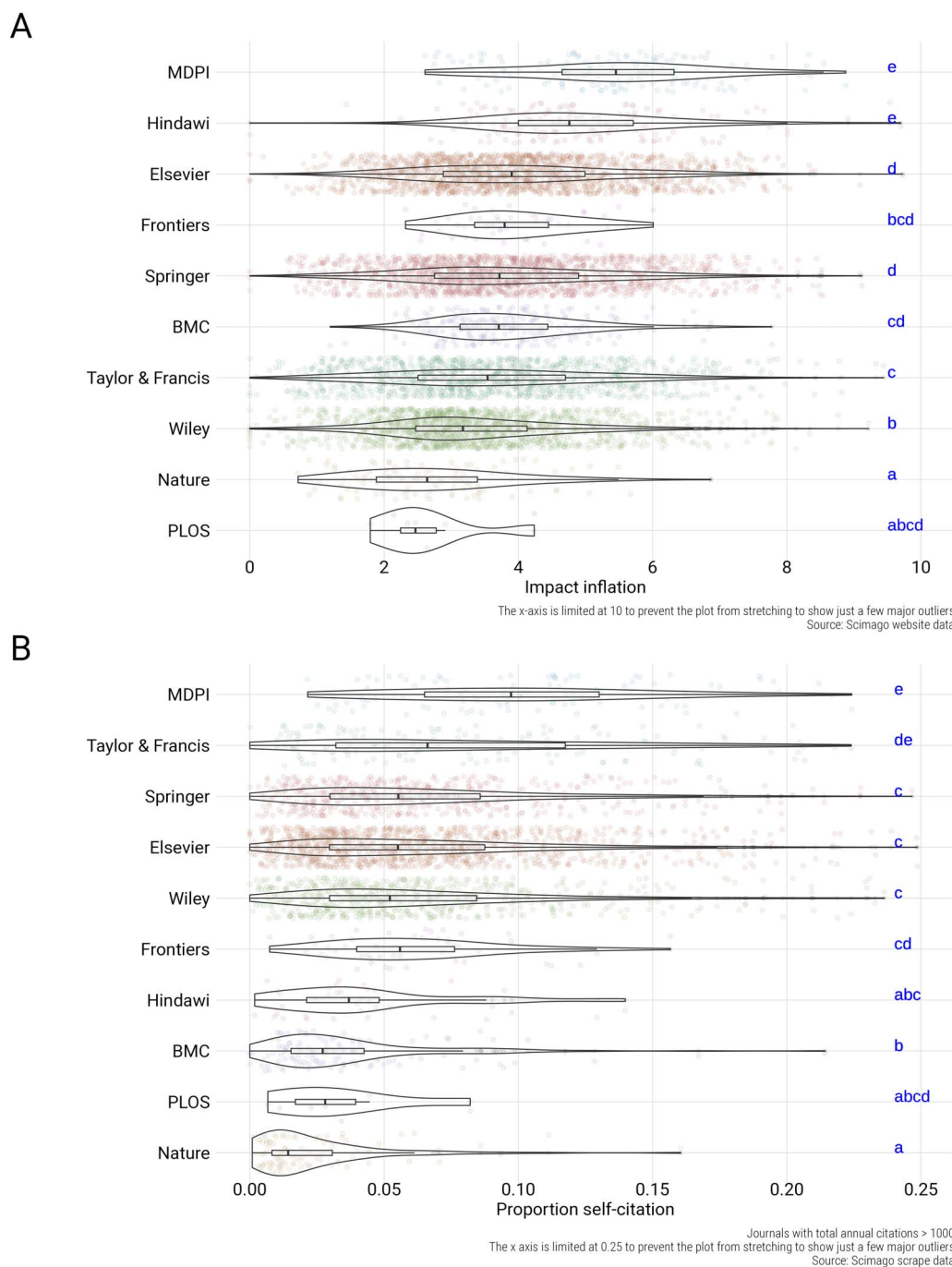


Figure 5. Changing behavior of citation metrics revealed by impact inflation (for journals with at least 1,000 annual citations). Statistical letter groups reflect significant differences ($p < .05$) in one-way ANOVA with Tukey HSD. A: MDPI and Hindawi have significantly higher impact inflation compared to all other publishers. Comparisons using samples of Clarivate IFs are shown in Figure S13. B: MDPI journals have the highest rate of within-journal self-citation among compared publishers, including in previous years (Figures S14 and S15). Here we specifically analyze journals receiving at least 1,000 citations per year to avoid comparing young or niche journals to larger ones expected to have diverse citation profiles.

publishers to earn prestige, and that IF can be abused by patterns of self-citation. We therefore asked if changes in journal citation behavior may have contributed to recent inflation of the IF metric.

To enable systematic analysis, we used Cites-per-doc from the Scimago database as a proxy of Clarivate IF (Cites-per-doc vs. IF: $R^2 = 0.77$, Figure S13). We then compared Cites-per-doc to the network-based metric SJR. Precise details of these metrics are discussed in the Supplementary methods (and see Guerrero-Bote & Moya-Anegón, 2012). A key difference between SJR and Cites-per-doc is that SJR has a maximum amount of “prestige” that can be earned from a single source. As such, if a journal disproportionately cites itself, or if citation cartels coordinate rings of citations between papers hosted by relatively few journals, this will increase the IF and Cites-per-doc of journals hosting such papers, but it will not increase the journal’s SJR. We define the ratio of Cites-per-doc to SJR (or IF to SJR) as “impact inflation.”

Impact inflation differs dramatically across publishers (Figure 5A), and has also increased across publishers over the last few years (Figure S14A). In 2022, impact inflation in MDPI and Hindawi were significantly higher than all other publishers ($p^{\text{adj}} < .05$). Interestingly, Frontiers had low impact inflation comparable to other publishers, despite growth patterns similar to MDPI and Hindawi.

The reason behind MDPI’s anomalous impact inflation appears to be straightforward: MDPI journals nearly universally spiked in rates of articles citing other articles from the same journal during the study period (here referred to as “self-citation” rate, Figure S14B).

There were significant differences in self-citation rates of MDPI journals compared to other publishers (Figure 5B, $p^{\text{adj}} < .05$ broadly, and MDPI vs. Taylor & Francis, $p^{\text{adj}} = .13$), including comparisons in previous years (S15, $p^{\text{adj}}_{2021} < .05$ broadly, including MDPI vs. Taylor & Francis $p^{\text{adj}}_{2021} = 3 \times 10^{-7}$). Indeed, beyond within-journal self-citations, in an analysis from 2021, MDPI journals received ~29% of their citations from other MDPI journals (MDPI, 2021), which would be rewarded per citation for IF but less so for SJR. Notably, Hindawi had self-citation rates more comparable to other publishers (Figures 5B and S15), despite high impact inflation. In this regard, while Hindawi journals may not directly cite themselves as often, they may receive many citations from a small network of journals, including many citations from MDPI journals (example in Figure S16), or from paper mills that operated within Hindawi (Bishop, 2023), and may have received biased citations from studies published via the same paper mills outside Hindawi.

In summary, we provide a novel metric, “impact inflation,” that uses publicly available data to assess journal citation behaviors. Impact inflation describes how proportionate a journal’s total citations are compared to a network-adjusted approach. In the case of MDPI, there was also a high prevalence of within-journal self-citation, consistent with reports by Oviedo-García (2021) and MDPI itself (MDPI, 2021). However, high impact inflation and self-citation are not strictly correlated with other metrics we have investigated.

5. DISCUSSION

We have characterized the strain on scientific publishing, as measured by the exponential rise of indexed articles and the pressure this creates for scientists. The collective addition of nearly one million articles per year over the last 6 years alone costs the research community immensely, both in writing and reviewing time and in fees and article processing charges (Aczel, Szaszi, & Holcombe, 2021; Shu et al., 2018). Further, given our strict focus on indexed

articles, not total articles, our data likely underestimate the true extent of the strain – *the problem is even worse than we describe*.

The strain we characterize is a complicated problem, generated by the interplay of different actors in the publishing market. Funders want to get the best return on their investment, while researchers want to prove they are a good investment. The rise in scientific article output is only possible with the participation of researchers, who act as authors, reviewers, and editors. Researchers do this because of the “publish or perish” imperative (Grimes, Bauch, & Ioannidis, 2018), which rewards individual researchers who publish as much as possible, forsaking quality for quantity. On the other hand, publishers host and spur the system’s growth in their drive to run a successful business (see File S1). Publishers structure the market and control journal reputation, and as such are focal players—leading to concerns regarding to what extent publisher behavior is motivated by profit (Butler et al., 2023; Ioannidis et al., 2023; Larivière et al., 2015). Indeed, a previous study found that publishers have used the growth of their corpus to negotiate for increases to journal subscription fees, with the result being that the value of money spent per citable article actually declined (Shu et al., 2018). Growth could be welcome if it comes from combatting systemic biases in academic publishing. However, if growth does not provide proportionate additional value, it increases strain on the scientific infrastructure.

Considering our metrics in combination (Table 3) also allows us to identify common trends and helps to characterize the role that different publishers play in generating this strain. Across publishers, article growth is the norm, with some groups contributing more than others. Impact factors and impact inflation have both increased universally, exposing the extent to which the publishing system itself has succumbed to Goodhart’s law. Nonetheless, the vast majority of growth in total indexed articles has come from just a few publishing houses following two broad models.

For older publishing houses (e.g., Elsevier, Springer), growth was not driven by major growth across all journals, but by the synergy of mild growth in both total journals and articles per journal in tandem (as discussed in Khelifaoui & Gingras, 2022). Another strategy used by certain for-profit gold open access publishers (e.g., MDPI, Frontiers, Hindawi) consisted of an increased use of special issue articles as a primary means of publishing. This trend was coupled with uniquely reduced turnaround times, and in specific cases, high impact inflation and reduced rejection rates. Despite their stark differences, the amount of strain generated through these two strategies is comparable.

The rich context provided by our metrics also provides unique insights. Ours is the first study, of which we are aware, to document that special issue articles are systematically handled differently from normal submissions: Special issues have lower rejection rates, and both lower and seemingly more homogeneous turnaround times. We also highlight the unique view one gets by considering different forms of citation metrics, and develop impact inflation (IF/SJR) as a litmus test for the degree of citation gaming taking place in a journal. Due to paper mill activity infiltrating the academic literature (Abalkina, 2023; Bishop, 2023), there is new-found interest in assessing the citation behaviors of publishing groups (Seeber, Cattaneo, & Birolini, 2024). Impact inflation addresses this using publicly available data, making it a valuable open science tool for assessing citation behaviors (in accord with recommendations by Waltman [2016]).

Throughout our study, MDPI was an outlier in every metric—often by wide margins. MDPI had the largest growth of indexed articles (+1,080%) and proportion of special issue articles (88%), shortest turnaround times (37 days), decreasing rejection rates (−8 percentage points), highest impact inflation (5.4), and the highest within-journal mean self-citation rate (9.5%).

Table 3. Strain indicators from 2016 to 2022. Data on total articles and impact inflation drawn from the Scimago data set. Data on special issue articles, turnaround times, and rejection rates come from web scrapes limited to the publishers shown. Rejection rate change for Elsevier and Hindawi starts from 2018 and 2020 respectively.

	2022					Change 2016–2022				
	Total articles (1,000s)	Share special issue (%)	Turnaround time (days)	Rejection rate (%)	Impact inflation	Total articles (%)	Share special issue (pp)	Turnaround time (days)	Rejection rate (pp)	Impact inflation
Overall	2,816	38	116	62	3.3	+47	+27	–23	–1	+1.1
Elsevier	498	–	134	71	4.0	+41	–	–4	+5	+1.5
MDPI	264	88	37	40	5.4	+1,080	+14	–28	–8	+2.2
Springer	251	3	157	–	3.9	+52	–1	+5	–	+1.5
Wiley	237	5	145	–	3.4	+36	–2	+5	–	+1.2
Frontiers	114	69	72	48	4.0	+675	+20	–25	+14	+1.8
Taylor & Francis	105	–	187	–	3.7	+59	–	+32	–	+1.5
Nature	57	11	185	–	2.8	+32	+6	+49	–	+1
BMC	44	10	162	–	3.9	+73	+1	+5	–	+1.5
Hindawi	39	62	83	74	5.0	+139	+36	–10	+3	+1.9
PLOS	19	1	198	59	2.6	–23	–3	+50	–4	+1.1

pp = “percentage points.”

Ours is not the first study analyzing MDPI (Abalkina, 2023; Copiello, 2019; Khoo, 2019; Oviedo-García, 2021), but our broader context highlights the uniqueness of their profile and their contribution to the strain.

Some metrics appear to be principally driven by publishers' policies: Rejection rates and turnaround time means and variances are largely independent of any other metric we assayed. This raises questions about the balance between publisher's oversight and scientific editorial independence. This balance is essential to maintain scientific integrity and authority: Oversight should be sufficient to ensure rigorous standards, but not so invasive as to override the independence of editors. Understanding how editorial independence is maintained in current publishing environments, though beyond the scope of this paper, is key to maintaining scientific integrity and authority.

Given the importance of scientific publishing, it is unfortunate that the basic data needed to inform an evidence-based discussion are so hard to collect. This discussion on academic publishing would be easier if the metrics we collected were more readily available—we had to web scrape to obtain many pieces of basic information. The availability of our metrics could be encouraged by groups such as COPE (2023), which publishes guides on principles of transparency. We would recommend transparency for proportion of articles published through special issues (or other collection headings), article turnaround times, and rejection rates. Rejection rates, in particular, would benefit from an authority providing a standardized reporting protocol, which would greatly boost the ability to draw meaningful information from them. While not a metric we analyzed, it also seems prudent for publishers to be transparent about revenue and operating costs, given that much of the funding that supports the science publishing system comes from taxpayer-funded or nonprofit entities. Referees such as Clarivate should also be more transparent; their decisions can have a significant impact on the quality of publisher stamps (see Table S2 and MDPI, 2023), and yet the reasoning behind these decisions is opaque.

Greater transparency will allow us to document the strain on scientific publishing more effectively. However, it will not answer the fundamental question: How should this strain be addressed? Addressing strain could take the form of grassroots efforts (e.g., researcher boycotts) or authority actions (e.g., funder or committee directives, index delistings). Researchers, though, are a disparate group and collective action is hard across multiple disciplines, countries, and institutions. Funders can more easily change the publish or perish dynamics for researchers, thus limiting their drive to supply articles. We recommend funders review the metrics we define here and adopt policies such as narrative CVs that highlight researchers' best work over total volume (DORA, 2022), which mitigate publish or perish pressures. Indeed, researchers agree that changes to research culture must be principally driven by funders (Wellcome Trust, 2020), whose financial power could also help promote engagement with commendable publishing practices.

Our study shows that regulating behaviors cannot be done at the level of publishing business model. Gold open access, for example, does not necessarily add to strain, as gold open access publishers like PLOS (not-for-profit) and BMC (for-profit) show relatively normal metrics across the board. Rather, our findings suggest that addressing strain requires that action be taken to address specific publishers and specific behaviors. For instance, collective action by the researcher community, or guidelines from funders or ethics committees, could encourage fewer articles to be published through special issues, which our study suggests are held to different standards from normal issues. Indeed, reducing special issue articles would already address the plurality of strain being added. Improved self-governance of publishers by the

publishing community, through, for example, more vigorous policing of new trends by institutions like COPE, might prevent the sorts of problems we have observed here from occurring.

Here we have characterized the strain on scientific publishing. We hope this analysis helps advance the conversation among publishers, researchers, and funders to reduce this strain and work towards a sustainable publishing infrastructure.

ACKNOWLEDGMENTS

We thank the following publishers for providing data openly, or upon request: MDPI, Hindawi, Frontiers, PLOS, Taylor & Francis, BMC, and The Royal Society. We further thank many colleagues and publishers for providing feedback on this manuscript prior to its public release: Matthias Egger, Howard Browman, Kent Anderson, Erik Postma, Yuko Ulrich, Paul Kersey, Gemma Derrick, Odile Hologne, Pierre Dupraz, Navin Ramankutty, and representatives from the publishers MDPI, Frontiers, PLOS, Springer, Wiley, and Taylor & Francis.

AUTHOR CONTRIBUTIONS

Mark A. Hanson: Conceptualization, Data curation, Formal analysis, Writing—original draft, Writing—review & editing. Pablo Gómez Barreiro: Conceptualization, Data curation, Formal analysis, Writing—review & editing. Paolo Crosetto: Conceptualization, Data curation, Formal analysis, Writing—review & editing. Dan Brockington: Conceptualization, Data curation, Writing—review & editing.

COMPETING INTERESTS

The authors have no competing interests.

FUNDING INFORMATION

This work was a labor of love, and was not externally funded.

DATA AVAILABILITY

Due to copyright concerns over our web scraping of information in the public domain, we were legally advised not to release our data and scripts publicly. In pursuit of principles of Open Science, we have shared what we can in the [Supplementary material](#) (including README files explaining to the reader how to assemble our Scimago data (File S2)). We have further developed a web app (via R shinyapp) hosted on the web page <https://the-strain-on-scientific-publishing.github.io/website/> that allows users to interrogate our data for specific trends. The app includes publishers not highlighted in the main text of this article. For more details see https://pagoba.shinyapps.io/strain_explorer/.

REFERENCES

- Abalkina, A. (2023). Publication and collaboration anomalies in academic papers originating from a paper mill: Evidence from a Russia-based paper mill. *Learned Publishing*, 36, 689–702. <https://doi.org/10.1002/leap.1574>
- Aczel, B., Szaszi, B., & Holcombe, A. O. (2021). A billion-dollar donation: Estimating the cost of researchers' time spent on peer review. *Research Integrity and Peer Review*, 6, 14. <https://doi.org/10.1186/s41073-021-00118-2>, PubMed: 34776003
- Bishop, D. V. M. (2023). Red flags for paper mills need to go beyond the level of individual articles: A case study of Hindawi special issues [Preprint]. *PsyArXiv*. <https://doi.org/10.31234/osf.io/6mbgv>
- Butler, L.-A., Matthias, L., Simard, M.-A., Mongeon, P., & Haustein, S. (2023). The oligopoly's shift to open access: How the big five academic publishers profit from article processing charges. *Quantitative Science Studies*, 4(4), 778–799. https://doi.org/10.1162/qss_a_00272
- Candal-Pedreira, C., Ross, J. S., Ruano-Ravina, A., Egilman, D. S., Fernández, E., & Pérez-Ríos, M. (2022). Retracted papers originating from paper mills: Cross sectional study. *BMJ*, 379, e071517. <https://doi.org/10.1136/bmj-2022-071517>, PubMed: 36442874

- COPE. (2023). *Best practices for guest edited collections*. Committee on Publication Ethics. <https://doi.org/10.24318/7cKLAia0>
- Copiello, S. (2019). On the skewness of journal self-citations and publisher self-citations: Cues for discussion from a case study. *Learned Publishing*, 32(3), 249–258. <https://doi.org/10.1002/leap.1235>
- CrossRef. (2023). [Data set]. <https://www.crossref.org/>
- DORA. (2022). *Changing the narrative: Considering common principles for the use of narrative CVs in grant evaluation*. <https://sfedora.org/2022/06/06/%EF%BF%BCchanging-the-narrative-considering-common-principles-for-the-use-of-narrative-cvs-in-grant-evaluation/>
- Else, H., & Van Noorden, R. (2021). The fight against fake-paper factories that churn out sham science. *Nature*, 591(7851), 516–519. <https://doi.org/10.1038/d41586-021-00733-5>, PubMed: 33758408
- Enserink, M. (2015). Open-access publisher sacks 31 editors amid fierce row over independence. *Science*, May 20. <https://doi.org/10.1126/science.aac4629>
- Fire, M., & Guestrin, C. (2019). Over-optimization of academic publishing metrics: Observing Goodhart's Law in action. *GigaScience*, 8(6), giz053. <https://doi.org/10.1093/gigascience/giz053>, PubMed: 31144712
- Fong, E. A., Patnayakuni, R., & Wilhite, A. W. (2023). Accommodating coercion: Authors, editors, and citations. *Research Policy*, 52(5), 104754. <https://doi.org/10.1016/j.respol.2023.104754>
- Fox, C. W., Albert, A. Y. K., & Vines, T. H. (2017). Recruitment of reviewers is becoming harder at some journals: A test of the influence of reviewer fatigue at six journals in ecology and evolution. *Research Integrity and Peer Review*, 2, 3. <https://doi.org/10.1186/s41073-017-0027-x>, PubMed: 29451533
- Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA*, 295(1), 90–93. <https://doi.org/10.1001/jama.295.1.90>, PubMed: 16391221
- Grainger, D. W. (2007). Peer review as professional responsibility: A quality control system only as good as the participants. *Biomaterials*, 28(34), 5199–5203. <https://doi.org/10.1016/j.biomaterials.2007.07.004>, PubMed: 17643484
- Grimes, D. R., Bauch, C. T., & Ioannidis, J. P. A. (2018). Modelling science trustworthiness under publish or perish pressure. *Royal Society Open Science*, 5(1), 171511. <https://doi.org/10.1098/rsos.171511>, PubMed: 29410855
- Grudniewicz, A., Moher, D., Cobey, K. D., Bryson, G. L., Cukier, S., ... Lalu, M. M. (2019). Predatory journals: No definition, no defence. *Nature*, 576(7786), 210–212. <https://doi.org/10.1038/d41586-019-03759-y>, PubMed: 31827288
- Guerrero-Bote, V. P., & Moya-Anegón, F. (2012). A further step forward in measuring journals' scientific prestige: The SJR2 indicator. *Journal of Informetrics*, 6(4), 674–688. <https://doi.org/10.1016/j.joi.2012.07.001>
- Hofseth, L. J. (2018). Getting rigorous with scientific rigor. *Carcinogenesis*, 39(1), 21–25. <https://doi.org/10.1093/carcin/bgx085>, PubMed: 28968787
- Ioannidis, J. P. A., Pezzullo, A. M., & Boccia, S. (2023). The rapid growth of mega-journals: Threats and opportunities. *JAMA*, 329(15), 1253–1254. <https://doi.org/10.1001/jama.2023.3212>, PubMed: 36939740
- Khelfaoui, M., & Gingras, Y. (2022). Expanding *Nature*: Product line and brand extensions of a scientific journal. *Learned Publishing*, 35(2), 187–197. <https://doi.org/10.1002/leap.1422>
- Khoo, S. Y.-S. (2019). Article processing charge hyperinflation and price insensitivity: An open access sequel to the serials crisis. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 29(1), 1–18. <https://doi.org/10.18352/lq.10280>
- Langin, K. (2023). *U.S. scientific leaders need to address structural racism, report urges* [Data set]. <https://doi.org/10.1126/science.adh1702>
- Larivière, V., Haustein, S., & Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLOS ONE*, 10(6), e0127502. <https://doi.org/10.1371/journal.pone.0127502>, PubMed: 26061978
- Liu, S.-N. C., Brown, S. E. V., & Sabat, I. E. (2019). Patching the “leaky pipeline”: Interventions for women of color faculty in STEM academia. *Archives of Scientific Psychology*, 7(1), 32–39. <https://doi.org/10.1037/arc0000062>
- Maher, D., Aseffa, A., Kay, S., & Tufet Bayona, M. (2020). External funding to strengthen capacity for research in low-income and middle-income countries: Exigence, excellence and equity. *BMJ Global Health*, 5(3), e002212. <https://doi.org/10.1136/bmjgh-2019-002212>, PubMed: 32206346
- MDPI. (2021). *Comment on: «Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI)» from Oviedo-García*. <https://www.mdpi.com/about/announcements/2979>
- MDPI. (2023). *Clarivate discontinues IJERPH and JRFM coverage in Web of Science*. <https://www.mdpi.com/about/announcements/5536>
- Meijaard, E., Cardillo, M., Meijaard, E. M., & Possingham, H. P. (2015). Geographic bias in citation rates of conservation research. *Conservation Biology*, 29(3), 920–925. <https://doi.org/10.1111/cobi.12489>, PubMed: 25817796
- Mlinarić, A., Horvat, M., & Šupak Smolčić, V. (2017). Dealing with the positive publication bias: Why you should really publish your negative results. *Biochemia Medica*, 27(3), 030201. <https://doi.org/10.11613/BM.2017.030201>, PubMed: 29180912
- Nakamura, G., Soares, B. E., Pillar, V. D., Diniz-Filho, J. A. F., & Duarte, L. (2023). Three pathways to better recognize the expertise of Global South researchers. *npj Biodiversity*, 2(1), 17. <https://doi.org/10.1038/s44185-023-00021-7>, PubMed: 39242772
- National Science Foundation. (2022). *Higher education in science and engineering* (HED-29; Numéro HED-29). National Science Foundation. <https://ncses.nsf.gov/pubs/nsb20223/figure/HED-29>
- Neff, B. D., & Olden, J. D. (2010). Not so fast: Inflation in impact factors contributes to apparent improvements in journal quality. *BioScience*, 60(6), 455–459. <https://doi.org/10.1525/bio.2010.60.6.9>
- Oviedo-García, M. Á. (2021). Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI). *Research Evaluation*, 30(3), 405–419a. <https://doi.org/10.1093/reseval/rvab020>
- Parolo, P. D. B., Pan, R. K., Ghosh, R., Huberman, B. A., Kaski, K., & Fortunato, S. (2015). Attention decay in science. *Journal of Informetrics*, 9(4), 734–745. <https://doi.org/10.1016/j.joi.2015.07.006>
- Peterson, C. J., Orticio, C., & Nugent, K. (2022). The challenge of recruiting peer reviewers from one medical journal's perspective. *Baylor University Medical Center Proceedings*, 35(3), 394–396. <https://doi.org/10.1080/08998280.2022.2035189>, PubMed: 35518802
- Quan, W., Chen, B., & Shu, F. (2017). Publish or impoverish: An investigation of the monetary reward system of science in China (1999–2016). *Aslib Journal of Information Management*, 69(5), 486–502. <https://doi.org/10.1108/AJIM-01-2017-0014>
- Sarewitz, D. (2016). The pressure to publish pushes down quality. *Nature*, 533(7602), 147. <https://doi.org/10.1038/533147a>, PubMed: 27172010

- Scimago. (2023). *SCImago journal & country rank* [Portal]. <https://www.scimagojr.com>
- Seeber, M., Cattaneo, M., & Birolini, S. (2024). Academic publishing business models: Self-citations and the selectivity-reputation trade-off. *SocArXiv*. <https://doi.org/10.31235/osf.io/5t8v7>
- Severin, A., & Chataway, J. (2021). Overburdening of peer reviewers: A multi-stakeholder perspective on causes and effects. *Learned Publishing*, 34(4), 537–546. <https://doi.org/10.1002/leap.1392>
- Shu, F., Mongeon, P., Haustein, S., Siler, K., Alperin, J., & Larivière, V. (2018). Is it such a big deal? On the cost of journal use in the digital era. *College & Research Libraries*, 79(6), 785–798. <https://doi.org/10.5860/crl.79.6.785>
- Thelwall, M., Kousha, K., Makita, M., Abdoli, M., Stuart, E., ... Levitt, J. (2023). In which fields do higher impact journals publish higher quality articles? *Scientometrics*, 128(7), 3915–3933. <https://doi.org/10.1007/s11192-023-04735-0>
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391. <https://doi.org/10.1016/j.joi.2016.02.007>
- Wellcome Trust. (2020). *What researchers think about the culture they work in* [Executive Summary]. <https://wellcome.org/sites/default/files/what-researchers-think-about-the-culture-they-work-in.pdf>
- Zwetsloot, R., Corrigan, J., Weinstein, E., Peterson, D., Gehlhaus, D., & Fedasiuk, R. (2021). *China is fast outpacing U.S. STEM PhD growth* [Data Brief]. <https://cset.georgetown.edu/wp-content/uploads/China-is-Fast-Outpacing-U.S.-STEM-PhD-Growth.pdf>