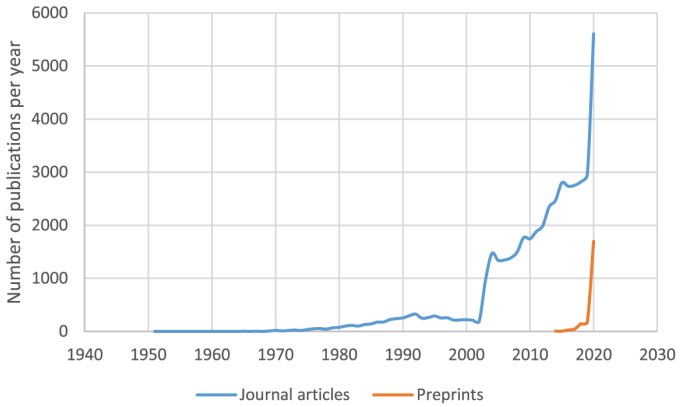
Biosecurity, Pandemic Preparedness and the three Ps of Open Science: Preprints, Peer-review and Preregistration

[work in progress]

Pre-Pandemic Preprints before the revolution

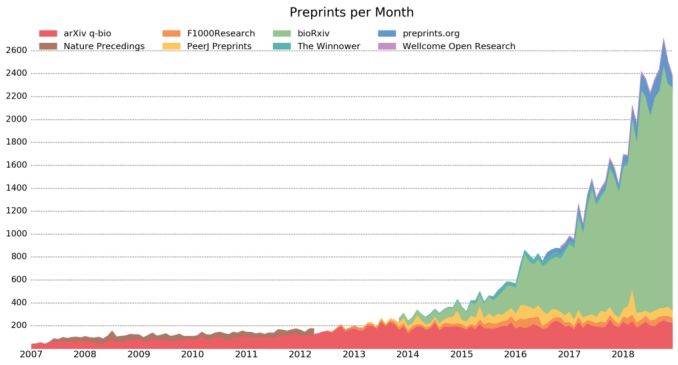
When COVID-19 struck, science needed to move at unprecedented speed. Enter the preprint -a research paper shared publicly before peer review, bypassing traditional publishing bottlenecks, including peer-review. In the life sciences, most preprints (between two-thirds and three-quarters) are concurrently submitted for peer-review and subsequently published in a journal, but some authors choose preprint deposition as the only way of communicating their research. Preprints provide a significantly faster way of publishing research than slogging through the system required by established journals; as such, they represent a radical departure from the traditional publishing model, and also serve as the foundation of an emerging 'publish, review then curate' model.

There are a number of reasons why authors would choose to use preprints. Firstly, because preprints forgo the long and often arduous peer-review process, preprints allow researchers to stake their claim on a discovery and establish priority. One of the main perceived risks of posting preprints has been the fear of getting 'scooped', where research is published before a rival researcher can publish theirs on the same topic, or where an idea or results are published without properly attributing credit to the researcher that came up with it first). Researchers have, however, become increasingly aware of the reality that instead, preprints likely reduce the chances of scooping because they allow authors to get their research out without the delays of journal publishing. Secondly, preprints have been demonstrated to increase visibility and citation counts of a research paper. Finally, preprints provide earlier feedback from a broader community of experts.



The number of journal articles and preprints per year, showing a sharp increase for both publication types since COVID-19 outbreak, with earlier pandemics (e.g. SARS, 2002) are also visible (Horbach, 2020)

The use of preprints is not new: the preprint server arXiv was launched in 1991 for physics, and SSRN, for social sciences, in 1994. While the biomedical sciences have been far slower to adopt preprints, with bioRxiv for biology launched in 2013 and medRxiv for medicine in 2018, preprints in these fields grew enormously in the half-decade preceding the pandemic, before experiencing an unprecedented surge during the COVID-19 pandemic.



Preprints per month on various preprint servers, 2007-2019 (Prepubmed, 2019)

Preprints during Ebola, Zika and COVID-19 outbreaks

Until early 2020, preprints had seen a steady increase in adoption, with previous outbreaks like Ebola and Zika prompting largely transient but significant increases, but from the earliest moments of the COVID-19 pandemic, they took on unprecedented proportions and significance. The rapid dissemination of knowledge facilitated by preprints enabled speedy updates to treatments and public health strategies and our reliance on them revealed the relative inflexibility of traditional peer-reviewed publication systems.

The push for preprints and transparent research

In the early months of the Ebola outbreak, researchers independent from official response efforts shared machine-readable surveillance reports and posted them on public repositories, while other teams aided efforts by rapidly depositing genetic sequences of the Ebola virus into public databases. As a result, a wider community of scientists were able to contribute key insights - 82% of epidemiological modelling publications on the epidemic relied solely on publicly available data.

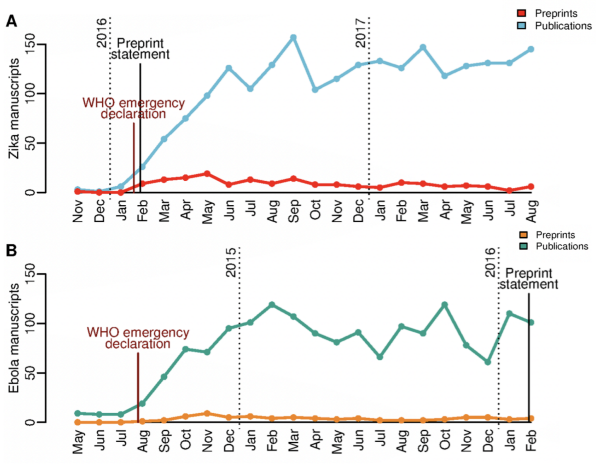
These developments came alongside calls from across the community to share data as openly as possible to ensure an effective response. Drawing lessons from the global response, towards the end of the epidemic an important WHO-held consultation in late 2015 affirmed the importance of timely and transparent pre-publication sharing of data during PHEs moving forward, noting that the exchange of critical data for vaccine development was faster than ever.. The scope of data sharing goes far beyond preprints, of course, but it's critical to note that while there was increasing discussion surrounding preprints at the time, I struggled to find any that drew associations between the value of preprints and the data sharing lessons from Ebola. That appears to have begun with the Zika epidemic.

Soon after the WHO declared a PHEIC on the Zika epidemic in February 2016, dozens of the world's most prestigious public health journals and funding agencies, led by the Wellcome Trust, issued a landmark statement, echoing the earlier WHO consultation during Ebola, on the importance of preprints and data sharing, stressing that the "the consequences of not doing so, have been thrown into stark relief by the Ebola and Zika outbreaks". Crucially, signatories agreed that "any data or preprint deposited for unrestricted dissemination ahead of submission of any paper will not pre-empt its publication in these journals". A few months later, academics came together to publish an article in science stating that "The time is right for biologists to post their research findings onto preprint servers" and in 2017, the NIH welcomed grant applications to cite preprints and encouraged their use to "to speed the dissemination and enhance the rigor of their work". What effect did all of these initiatives have on preprint adoption?

Preprints respond to the call

It's worth highlighting that while calls made by official research bodies for the use of preprints undoubtedly would have spurred on their adoption, the significant rise of preprints between Ebola, Zika and COVID-19 also represent a bottom-up, decentralised response among scientists. Many virologists, epidemiologists and other researchers may have drawn directly from the experience of previous outbreaks, tipping the balance in favour of adopting preprints, often for the first time.

Let's look at the statistics on preprints posted during the Ebola and Zika epidemics. There were roughly 75 Ebola preprints and 174 Zika preprints (search periods *West African Ebola Virus, 05/2014 -01/2016; Zika Virus, 11/ 2015 -08/2017* - similar durations for comparison) with the respective virus name in the title or abstract. Small numbers in absolute terms, but a 132% increase in the space of 1-2 years. This is compared to only a 33% increase in peer-reviewed journal articles across the same periods and search criteria. Just as important, however, is that among the subset of preprints that were matched with an eventual publication, the proportion with original data increased significantly from 7% for Ebola to 46% for Zika.

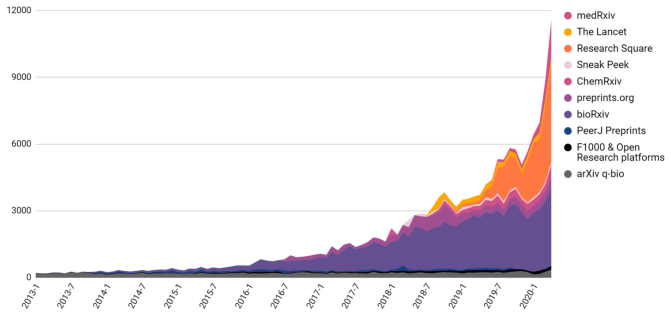


Total number of newly posted preprints (A: Zika - preprints in red, peer-reviewed publications in blue; B: Ebola- preprints in orange, peer-reviewed ublications in green) by month (Johansson et al. 2018). Black line for 'Preprint statement' refers to Wellcome Trust statement

Nevertheless, despite the 2016 statement and the broad acknowledgement of the need for preprints, less than 5% of the journal articles about the two epidemics were posted first as a preprint.

The floodgates open

The WHO declared a PHEIC on COVID-19 on 30th January 2020. The very next day, the Wellcome Trust issued a call for changing business as usual that urged authors to put up COVID-19 manuscripts as preprints. Let's look closer at the unprecedented contribution of preprints in the early months of the pandemic response.

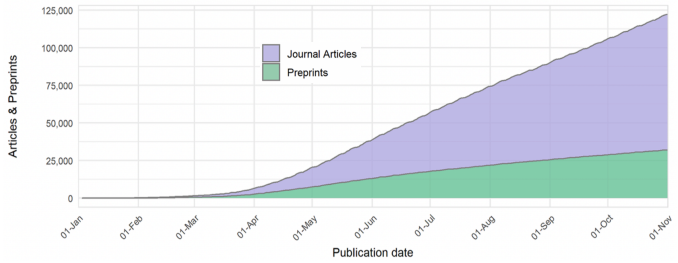


Biomedical preprints per month, by source, from 2013 - April 2020 (Polka and Penfold, 2020)

In the first four months since the PHEIC, a staggering 19,389 peer-reviewed articles about COVID-19 were published, over around 35% which were preprints posted on public preprint servers. By this time, more than half of biochemistry, biophysics, and genomics preprints became published articles, compared with 29% in epidemiology and 26% in bioinformatics. As a result, the proportion of first publications as preprints vs in journals would be even higher. It should be noted that many journals still don't allow for a manuscript to be published as a preprint before journal submission, restricting authors' choice for journal submission as

well as preventing manuscripts from being posted as preprints. During the same time window a year earlier, in 2019, only 3% of the biomedical literature were preprints. And compared to previous outbreaks, that's up from 4.4% and 7.4% of the total outbreak-related publications during Ebola and Zika outbreaks, respectively. bioRxiv and medRxiv alone hosted 25% of all COVID-19 related scientific research published in the first 10 months of the pandemic, amounting to 10,000 papers on these sites.

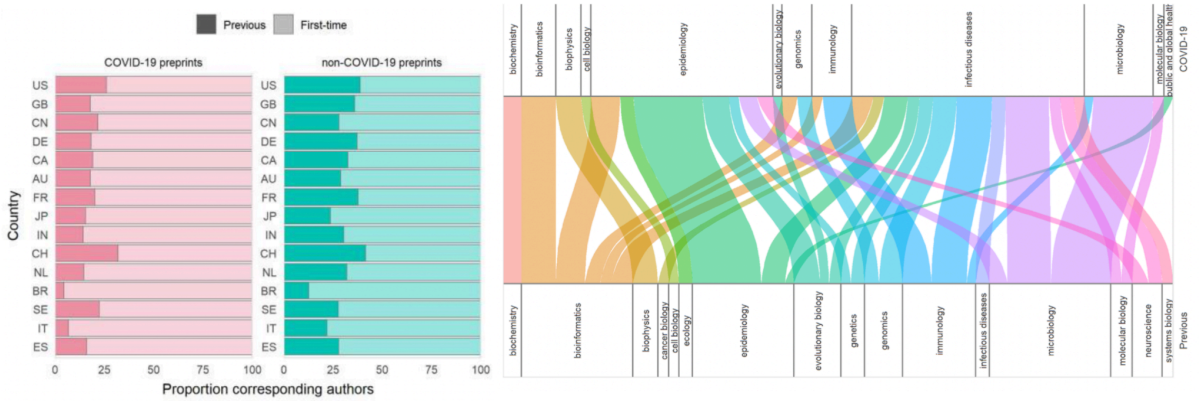
As the pandemic progressed, a number of further projects and initiatives were launched to encourage preprint use and facilitate a rapid review system. In May 2020, NIH began a preprint initiative to index preprints that come from NIH-funded projects and increase the discoverability of early NIH research posted on preprint servers by making them available on PubMed. In June, a new COVID-19 journal was launched, *Rapid Reviews: COVID-19*, which, publishes preprint articles after a thorough peer-review. The Sinai Immunology Review Project worked to review and validate COVID-19 related preprints posted to medRxiv and bioRxiv servers. The Review Commons platform, launched in December 2019, is another initiative on that direction, where preprint authors are offered an opportunity to request a journal-independent portable peer review [31]. Lastly, it is worth mentioning new editorial policies from *eLife* that resulted in the launch of Preprint Review [32] and made the preprint deposition mandatory prior to a journal submission [33].



Cumulative growth of journal articles and preprints containing COVID-19–related search terms (Fraser et al., 2021)

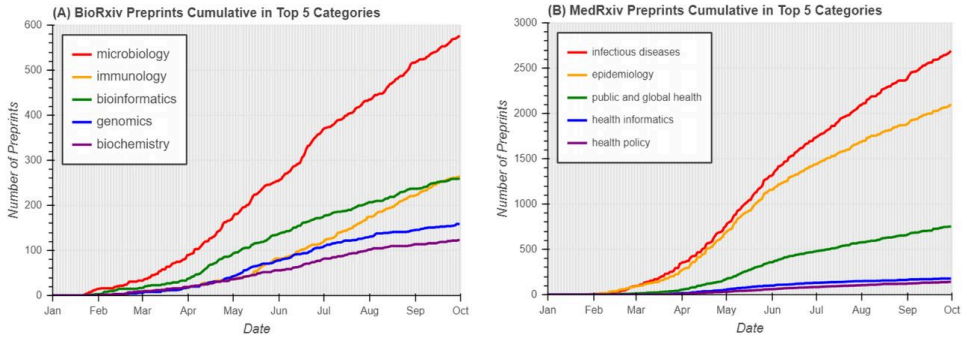
With the pandemic, a new category of preprint authors emerged from around the world, publishing across and between fields. Among COVID-19 authors, 85% were posting a preprint for the first time, compared to 69% of non-COVID-19 authors in the same period. The largest proportion of preprints were posted by authors in the US, followed by the UK and China. Notably, China is overrepresented in terms of COVID-19 preprints relative to its non-COVID-19 preprint output (*figure below, left)*.

During the pandemic, authors posted preprints extensively on fields spanning immunology, microbiology, infectious diseases, and epidemiology. There were concerns surrounding the appropriateness of researchers from other fields switching to COVID-19 research to capitalise on the attention. The incentives are such that this would undoubtedly have occurred to some degree. Reassuringly, however, most authors - comparing their COVID-19 preprint with their previous preprint - were not drastically changing fields, with category differences generally remaining stable or moving between related areas (*figure below, right*).



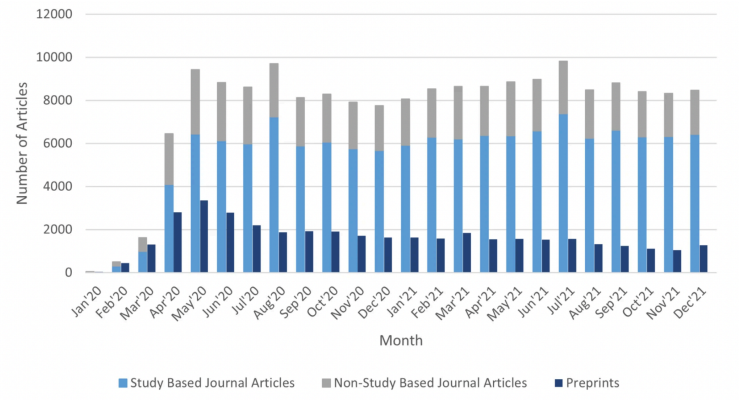
(Left) Proportions of COVID-19 and non-COVID-19 corresponding authors from each of the top 15 countries; (Right) Change in bioRxiv/medRxiv preprint posting category for COVID-19 preprint authors compared to their previous COVID-19/non-COVID-19 preprint (Fraser et al., 2021)

Researchers likely turned to preprints during the pandemic, sometimes for the first time, as a means of rapid and transparent scientific communication rather than, say, seeking pre-submission feedback, given that preprints were typically submitted concurrently with peer-reviewed journal submissions. Only a third of COVID-19 preprints posted during the first nine months would later appear as peer-reviewed journal articles. This figure is lower compared to previous outbreaks, with 60% for Ebola, and 48% for Zika. As such, it's likely that some proportion of these COVID-19 preprints consisted of immediate and work-in-progress results the authors felt were unsuitable for peer-review, but nonetheless deemed important enough to disseminate anyway. This rise in preprints in the first few months of the pandemic is especially staggering in microbiology and bioinformatics (on BioRxiv) and infectious diseases and epidemiology (on MedRxiv).



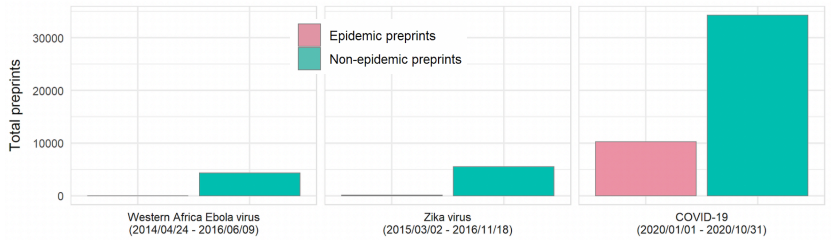
Cumulative COVID-19 preprint submissions in the top 5 categories to BioRxiv and MedRxiv in the first months of the pandemic (Sevryugina and Dicks, 2021)

In general, the relative reliance on preprints as a mode of disseminating research is particularly pronounced during the early months of the pandemic, a trend that mirrors the preprint response to the two previous outbreaks, albeit on a far larger scale. Depending on the source, some use the submission timestamp of eventual publications while others use the publication date. Given the long delays in publishing (more on that later), using the publication date will obviously skew the journal article figures to later in the year, which can give the false impression of a higher-than-actual reliance on preprints, especially in the first few months of the pandemic response.



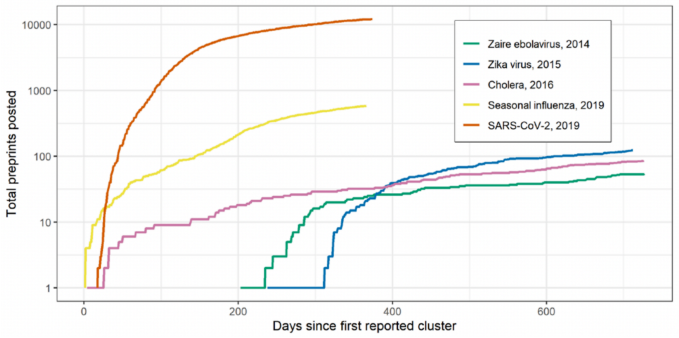
COVID-19-related preprints and journal articles published, by month, January 2020 to December 2021. 'Study Based Journal Articles': articles that analyzed data such as observational studies, trials, and meta-analyses. 'Non-Study Based Journal Articles': articles that did not analyze data such as commentaries (Drzymalla et al., 2022)

Critically, the surge in outbreak-related preprints during the COVID-19 pandemic is not explained by the broader trend of increasing preprint server usage: pandemic-related preprints represented a far larger proportion of total preprints during the COVID-19 pandemic compared to the Ebola or Zika outbreaks.



Total number of epidemic and non-epidemic related preprints posted on bioRxiv and medRxiv during multiple epidemics: Western Africa Ebola virus, Zika virus, and COVID-19 (Fraser et al., 2021)

An early response to pandemics and biosecurity threats is critical, and in addition to the scale, the preprint response to the COVID-19 pandemic was earlier and hugely accelerated, increasing at rates well above 100 times greater, compared to previous outbreaks.



"Cumulative frequency curves of preprints addressing five pathogens causing epidemics within the last five years, aggregating across three preprint servers (bioRxiv, medRxiv, arXiv), time-adjusted to represent days since first official health authority notification of a case cluster" (Brierley, 2021)

Overall, the reliance on preprints played a crucial role in outbreak responses. During the COVID-19 pandemic, the research community embraced preprints more than in previous outbreaks, leading to (1) a significantly higher number of preprints overall; (2) a greater proportion of epidemic-related publications released as preprints; (3) a larger share of preprints specifically focused on the epidemic; (4) faster publication of epidemic-related preprints at accelerated speeds

Preprints are crucial from a pandemic-preparedness and biosecurity perspective because they enable the rapid dissemination of research, ensuring that critical findings reach policymakers and healthcare professionals without the delays of traditional peer review. Biosecurity experts should be aware of the growing reliance on preprints during public health crises to effectively anticipate the opportunities and risks that could arise from their adoption.

A failed revolution?

Four years have passed since these events, and a sobering dose of nuance is needed here. As seen in past outbreaks, the initial surge in preprints stabilized after six months before fading into the background. While the scale of preprint growth was unprecedented, so too was the COVID-19 pandemic itself. Many open science advocates had heralded for a sweeping transformation in publishing that would last beyond the pandemic, but this did not materialize.

Despite enthusiasm for open science and the many initiatives for preprints, a 2021 RORI report found that only 5% of peer-reviewed COVID-19 studies had a corresponding preprint. The State of Open Data report also showed that while global awareness of data sharing has grown, authors' willingness to share has declined. A Springer Nature survey found that although 85% of authors support early sharing, opt-in rates remain low due to concerns over copyright, misuse of open data, and lack of recognition and credit.

Some open science advocates had heralded a revolution in publishing practices, but momentum for open science, at least on the preprint front, has slowed. While medRxiv saw a surge in COVID-19 preprints early in the pandemic, this declined by mid-2020. None of the critical phase 3 COVID-19 vaccine trials, supported by Oxford-AstraZeneca, Moderna, or Pfizer, were published as preprints . A 2022 Wellcome review found fewer than half of signatories to its data-sharing commitment provided clear access to COVID-19 data, revealing persistent transparency issues, especially in clinical trials.

Despite these shortcomings, the rapid expansion of open research practices during COVID-19 demonstrated the feasibility of radical, large-scale change. In the meantime, general disillusionment among researchers with traditional publishing shows little signs of letting up and the debates around open science continue to widen. In future pandemics and public health emergencies, a wholesale revolution in scientific publishing remains a possible, even likely scenario - and one biosecurity efforts should anticipate, and even capitalise on.

Traditional peer-review during outbreaks

The COVID-19 pandemic was the first global health crisis faced by the modern scientific publishing industry. The industry did not exist in any recognisable form at the time of the 1918 Influenza pandemic. Beyond its devastating death toll, the COVID-19 pandemic triggered immense economic and social crises, including the worst global recession since the Great Depression, widespread food shortages, and intensified political tensions. The COVID-19 pandemic is also the first to unfold in a digital age of internet publishing, social media and alternative, open forms of scientific dissemination. Under such conditions, how did the publishing industry respond to this crisis?

Aside from the long running debate over the idea of traditional peer review as the 'least-worst' filtering method for research quality and impact, critics often point to its slowness and inefficiency. Is traditional peer-review sufficiently responsive to the demands of an effective pandemic response? A clear assessment of peer review delays will help inform prioritisation strategies for biosecurity-relevant interventions at the publication stage.

For clarity, I'll use terms like peer review, review, and submission-acceptance interchangeably (in practice, articles undergo multiple rounds of peer review, revisions, and editorial feedback before acceptance).

How long does peer-review and publication take, actually?

Before we look at how publishing has responded to various outbreaks, let's first get a handle on how long it typically takes to publish scientific papers in 'peacetime'. It turns out that this is a frustratingly difficult question to answer. There are only a handful of sources on this question, and the lack of solid research on this will become a disturbingly recurrent theme throughout this piece.

A computational-biology graduate student at UCSF looked at all the papers in PubMed with listed submission and acceptance rates finding that the total peer-review process (submission-acceptance) has stayed steady at around 100 days for over 30 years, while production delays (acceptance to publication), driven primarily by the arrival and adoption of online publication in the 90s, have been slashed in half since the early 2000s to around 20-25 days.

However, the submission-acceptance delays here are likely an underestimate due to increasing "reject and resubmit" decisions by journals rather than "revise" to shorten their reported acceptance delays and the fact that many journals - including the most prestigious ones - don't always deposit time stamp data in PubMed, so there's the potential for selection bias. The same analysis indicated that on average, higher impact, and therefore more prestigious, journals take longer and that in many such journals, the median review time has actually gone up significantly. In the past decade, *Nature*'s median review time went from 85 days to above 150 days while at *PLoS ONE* it went from 37 to 125 days.

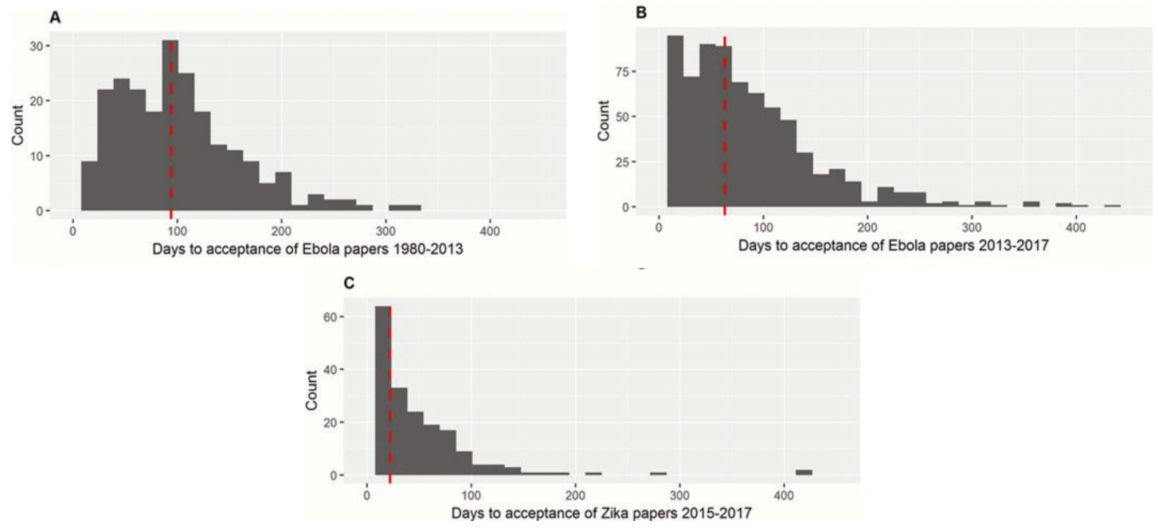
Note that these are submission to acceptance rates, so naturally, they only include papers that were eventually accepted. They don't represent the full delay between intention to

submit and publication because they exclude the time “shopping around” for a journal to publish their findings in, which often includes multiple rejections (before or after peer-review). In an attempt to calculate the delays of the full publication cycle, a cell biologist found that, assessing all the papers on which he was an author, the median time from *first* submission to being published online was **250 days** (just 2 weeks short of the average time it takes to have a baby!). The median reviewing time was 121 days and the “shopping around” phase ranged from a few days to more than eight months. More information on peer-review and total publication time can be found here. In the interests of not dragging this on further, I think we can definitively conclude that at least for the foreseeable future, peer-review and standard publication will continue to take… a very long time, in pandemic response terms. On the order of months to a year. Considerably longer than the few days, at most, after running some ‘sanity checks’, it can take to get a preprint published.

Previous outbreaks

Before we assess how we fared during covid, let’s again take Zika and Ebola as comparisons. Of the 5% of the journal articles about the two epidemics that were first posted as a preprint, most of them didn’t appear in a journal for another 100 days. However, while there wasn’t a major concerted, coordinated and public effort by journals to speed up peer-review on par with that of the COVID-19 pandemic, previous epidemics did see a decrease in publication times from the norm.

Ebola-related papers from 1980-2013 took a median of 94 days to acceptance compared to 63 days during the Ebola epidemic (a wide date range here, from 2013-2017 - during which time papers on herpes, a non-epidemic disease at the time, remained at 93 days) and 22 days for Zika related papers. Journals did, in fact, speed up peer-review times of outbreak-relevant papers during their respective outbreaks.



Days to acceptance of (A) Ebola papers 1980-2013, where Ebola emerged periodically in African countries (n=230) (B) Ebola papers during the West African Ebola epidemic 2013 - 2017, the largest outbreak to date by far (n=815) © Zika papers during the Zika epidemic 2015-2017 (n =248). Stipulated red line indicates median. (Ronit, 2017)

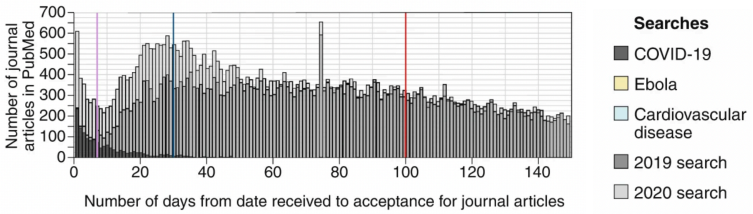
Picking up the pace of peer-review

In 2020, the speed of publication could not keep pace with the speed of research—or of the rapidly worsening pandemic. Scholarly journals and their publishers promised to speed up the publication of COVID-19 manuscripts. In particular, in early 2020, some publishers coordinated and launched the 'COVID-19 Rapid Reviewer Initiative', a collaborative initiative to generate a shared pool of rapid reviewers specifically for reviewing COVID-19 relevant articles. Prestigious journals like PLOS and some owned by Wiley, Elsevier, Sage and PeerJ implemented specific policies to fast-track COVID-19 papers. Peer-reviewed journals were and still remain a staple of knowledge dissemination during outbreaks, so their pledge to speed up the review process, the primary obstacle to research dissemination after findings are made, would be helpful to our ability to respond to the global crisis. Much was made about the successes, challenges and risks of speeding up peer-review, so let's take a look at how successful these efforts were.

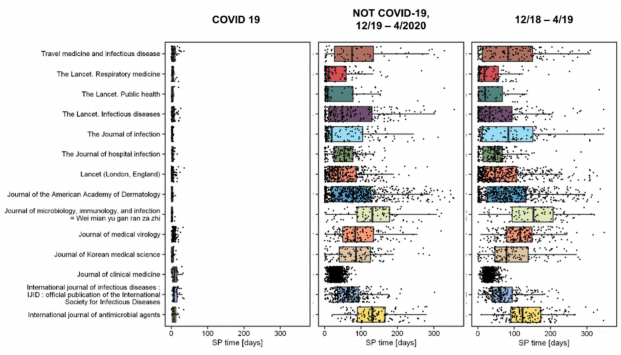
How effectively did publishers speed up peer-review and publication delays, actually?

There are major discrepancies in the figures arrived at in different studies as a result of an important, yet unmentioned (for whatever reason) statistical phenomenon. Concerningly, this fact attracted little attention at the time, so it's worth going through the major studies published in the first months of the pandemic.

A widely cited early 2020 study published in *Nature* found that COVID-19-related manuscripts between January 2020 and April 2020 had a median reviewing time of only 6 days (!) and that 59% of COVID-19 articles were published within a week, compared to only 3% for cardiovascular disease in 2019. This COVID-19-specific reviewing time is less than half of that for Ebola articles (in the first months of the West African Ebola epidemic) at 15 days.



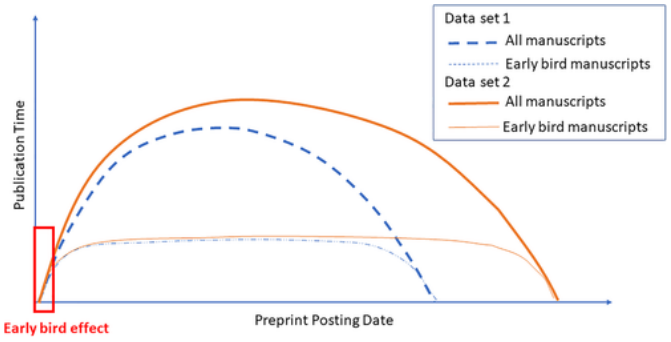
Another early 2020 study found that the total *submission-to-publication* time of COVID-19 articles in most journals reduced dramatically, with the times being reduced by a factor of 10-15 to a median of 5 days (!).



In the limitations section of this study, the authors noted "*Our methods are obviously limited by time point of the analysis, and we believe a repeated analysis of the available data could yield different results*".

I reference these two studies only because, indeed, later analyses would show that these figures were dramatically miscalculated. More precisely, they were calculated without even attempting to account for the scale of an intuitively obvious phenomenon (which they allude to here) that makes their figures compromised to the extent that they are (and should have been) unusable. Beyond this one line in one of the studies, no further acknowledgement was made or caution advised. These figures were published in prestigious journals (*Nature* and *Springer Nature*) and were widely quoted [source7] in articles in the first year of the pandemic, praising the dramatic and unprecedented success of rapid review times by publishers. **As such, they serve as important case studies for future research analysing publication times in response to sudden global events and emergencies.**

So what's the real figure, and what was the issue? Later studies can be found here and here which came up with rather different numbers, depending on the time frame and COVID-19 search criteria (57 days for COVID-19 articles, or just under 100 days for all articles as the months dragged on). Long story short, the shockingly low figures of 6 and 5 days reported in the early studies were so due to a simple 'early bird' effect described in a 2022 study.

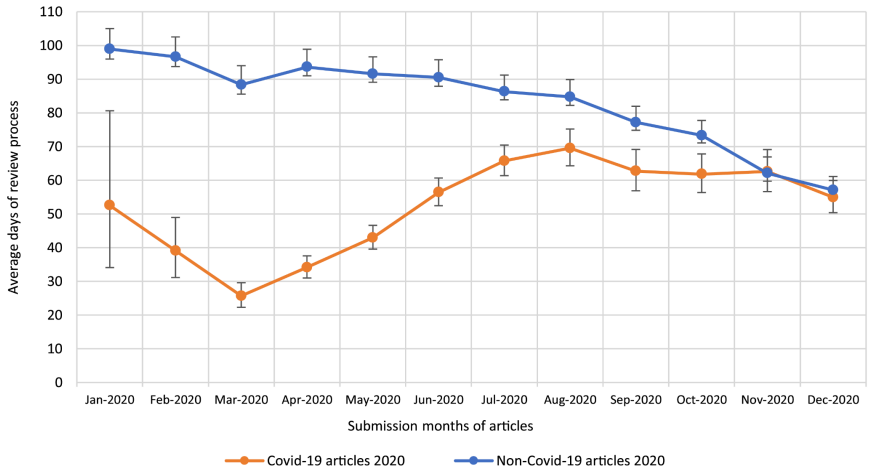


Basic schematic of 'early bird' effect in analysing publication times of manuscripts after newly emerging event (Sevryugina and Dicks, 2022)

The sudden onset of the COVID-19 pandemic meant that before January 2020, no COVID-19 publications existed. As a result, ‘early bird’ manuscripts (manuscripts that are published rapidly in response to an emerging event), normally continuously compensated by newly published papers with average or long publication delays, became the dominant manuscripts in the early data sets. The reported extremely short publication delays for manuscripts on ‘hot’ topics at the outset of a fast emerging event produced an exaggerated perception of a rush in publishing during the early period of the pandemic.

Running an updated analysis in 2022, the study found that publication delays for COVID-19 papers show only a moderate expediting as compared to pre-pandemic papers. Median peer-review and production stage delays were **66** and **15 days**, respectively (January 2020 to 31 March 2021).

The significant, but moderate, reduction in publication times was also relatively short-lived. Another 2022 study took the review speeds by **submission date**, finding that while review speeds of Covid-19 papers was especially rapid during the first 5 months (Jan-May 2020) of the pandemic, at 1.9–3.4 times faster than pre-pandemic, this speed advantage was no longer present for those submitted from November onwards. Overall, Covid-19 articles submitted during 2020 were reviewed 1.7–2.1 times faster than non-Covid-19 articles submitted 2017-2020, although heavily skewed by the first 6 months.



“Average days of review processes for Covid-19 articles (n = 4,038) and non-Covid-19 articles (n = 26,104) submitted to 100 medical journals with high impact factors between January and December 2020” (Kousha and Thelwall, 2022)

The study also investigated this on a per-journal basis, finding that the difference during and before the pandemic was statistically significant across 84% of journals. This is important because high-impact journals may generate bespoke pathways, whereas others may not have the resources for this. The findings also show that, across almost all journals, faster-published COVID-19 papers tended to garner a higher citation impact, although this is of course a complex correlation to parse for causation.

The limits and insights of peer-review during COVID-19

To be clear, this achievement is still incredibly impressive, and absolutely commendable. And it's to be expected that slashing industry-wide publication delays by half would not be sustainable long-term. A rapid response to the pandemic in research outputs is critical to inform healthcare and public policy in its early stages, so the transient speeding up of peer-review early on should be seen as a major success. The efforts to coordinate this likely saved an untold number of lives that would have been lost if we had been left to rely on our publication system in its default state.

Given the lack of transparency in peer-review and editorial processes, concrete details on how journals accelerated their workflows are scarce. However, we do know they created expedited pathways for COVID-19 submissions, assembled dedicated editorial teams, relaxed non-critical formatting requirements, and, perhaps most importantly, invited more reviewers per paper [source9]. This shift offers a rare opportunity to analyse what worked and what didn't in a system under unprecedented stress. Such insights could inform broader review processes now, in preparation for future emergency scenarios, and improve or further accelerate peer-review during future crises, significantly impacting our ability to respond to biosecurity threats.

Still, some feel that the events of 2020 was further proof that the traditional peer-review system is ill-suited to responding to societal needs. The rapid evolution of the pandemic situation - such as the sequential emergence of new variants - meant that even a month-long review process, remarkably fast under normal circumstances, felt unacceptably slow in a crisis where lives were lost daily. Complicating matters, the scientific publishing industry arguably exists as a vital component of a system that itself perpetuates the "publish or perish" culture that contributed to a flood of COVID-19 manuscripts - many with methodological flaws and low reproducibility.

[tbc]

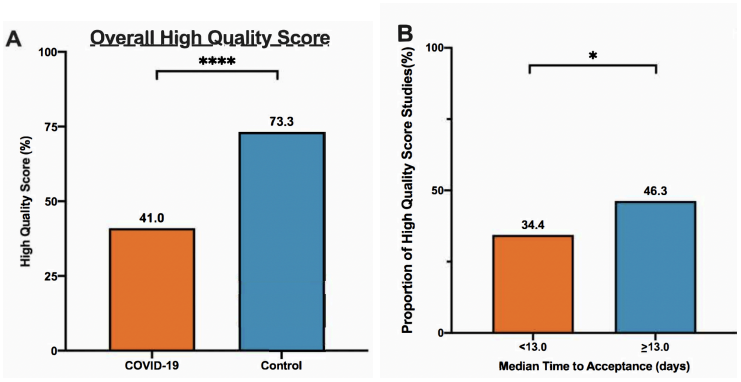
Concerns over quality

In order to achieve this feat, were COVID-19 related papers held to the same standards and scrutiny as non-COVID-19 papers during review? What proportion of lower quality papers were allowed to slip through the cracks? It's well known that during the pandemic, editors were far more lenient in accepting a wider range of study designs (case reports, smaller uncontrolled studies) than they normally would and that in the rush to publish, authors would sometimes be asked state the limitations of their data rather than amend, remove or recollect their poor quality data [source13]. Small sample sizes and low study power, in particular, are often cited as key contributors to the 'replication crisis' across all fields of research.

Of course, it's difficult to draw concrete conclusions about the direct effect of speeding up the review process on overall peer-reviewed paper quality because of selection bias. Were articles with stronger claims - and therefore potentially less nuance, caution and methodological rigour - more likely to be fast-tracked at a higher priority by editors? Or are

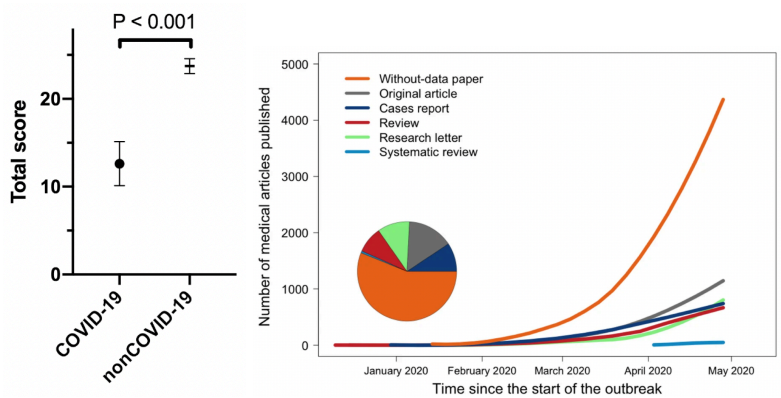
such studies inherently easier and quicker to review? Or was it, indeed, through cutting corners in the peer-review process, and thereby lowering effective acceptance standards, that led to the flood of poor quality papers passing peer-review?

On the lower quality of COVID-19 studies, studies were unanimous, although attributing the lower quality to any specific variable is challenging. One study showed that the methodological quality score of peer-reviewed COVID-19 articles was significantly lower compared to historical controls (figure below, with further insights in study). Interestingly, the same study also found that when stratifying COVID-19 articles by submission-acceptance time (above or below median of 13 days), decreased acceptance time was associated with lower methodological quality scores (specifically, lower scores for case series and cohort study designs but not for case–control nor diagnostic studies).



A: Lower methodological quality score associated with COVID-19 articles over control articles B: Lower quality case series studies score associated with COVID-19 articles over control articles

Another study concluded "that the quality of COVID-19 papers in the three highest ranked scientific medical journals is below the quality average of these journals" in their level of evidence and quantitative quality scores (figure below), as well as increasingly long authorships in lower quality COVID-19 publications. Another study also demonstrated that from the earliest weeks of the pandemic, "among peer-reviewed medical studies, publications without original data dominated the exponential growth of COVID-19 literature" (figure below). Similar evidence, particularly in the early phases of the pandemic, of low-quality COVID-19 research has been found among systematic reviews as well as a high level of spin in abstracts of randomised controlled trials.



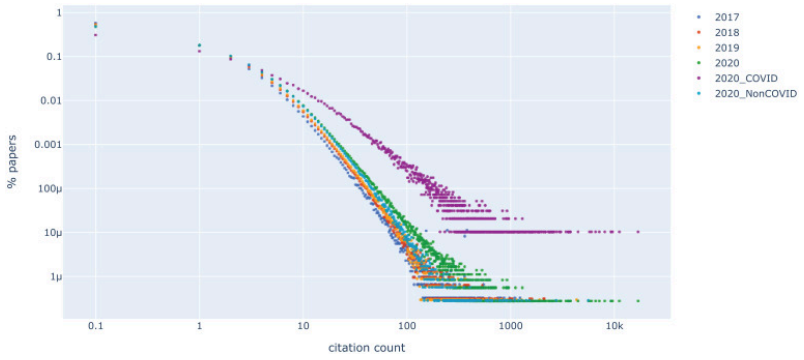
Left: Primary analysis for mean total quality scores of COVID-19 vs non-COVID-19 papers (Zdravkovic et al., 2020); Right: Study publication categories for COVID-19 papers over time (Raynaud et al., 2021)

It's also possible that the handling of potential conflicts of interest in the editorial process could also deteriorate under high-pressure, and that this could contribute to a lowering of peer-reviewed research quality. An investigation into the relationship between conflicts of interest and acceptance times found that among COVID-19 research articles, the frequency of editorial conflicts of interests increases with the decrease in acceptance times, from 18.6% for those articles accepted within 20 days, to 31.7% for those accepted in a day or less. A high-profile example of this is the infamous paper published by Didier Raoult and his research team, a paper which lies at the heart of the hydroxychloroquine saga (one of the biggest scientific controversies in the early months of the pandemic, which is explored in the case studies section). The editor-in-chief of the Q1 Elsevier-owned journal that published the paper also *happens* to work in Raoult's institute and reports to Raoult - he's also a signatory on the paper. The peer-review of this paper was abnormally fast, expedited in a single day. This presents obvious conflicts of interest that, although could have been mitigated, weren't.

Little research is available on how often conflicts of interest of this nature occur among peer-reviewed journals. Based on this phenomenon, this paper set out to calculate the proportion of contributions published in a journal by any single author as a rough index to spot problematic journals. Looking at infectious disease journals between 2015-2019, the 95th percentile of journals stood at around 10% (i.e. for 95% of journals, their most prolific author contributed to ≤10% of the publications) and journals above this threshold were analysed further. While the Elsevier-owned journal was a clear outlier, there were a number of others that also passed the threshold, and this indicator could be used by publishers (and the wider community) to vet and audit high-risk journals and investigate the editorial and peer-review processes in these journals. Overall, these findings raise concerns about the fairness and transparency of the editorial and peer-review process, especially for those papers with shorter acceptance times.

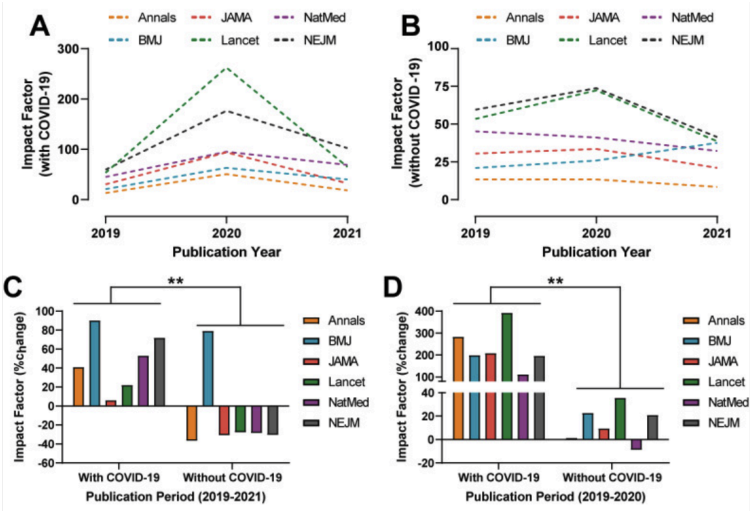
Too much too fast and a new 'citation elite'

Conflicts of interest aside, what's really driving this? During the COVID-19 pandemic, a significant distortion occurred in scientific publishing, driven by the incentive structures around journal impact factors. Though COVID-19 papers represented only 4% of scientific literature, they garnered a disproportionate 20% of all citations. In medical literature specifically, COVID-19 articles (17% of publications) received an astounding 80% of citations. This citation advantage was dramatic – COVID-19 papers published in 2020 received over 8 times more citations than non-COVID papers generally, and over 20 times more in medical literature. The incentives in the decision-making process for authors choosing to publish on COVID-19 are clear, too: while the total number of authors who received ≥100 citations during the pandemic was almost double for all non–COVID-19 work than for COVID-19 work, the difference was eliminated at ≥1,000 citations and >5,000-10,000 citations were received only for COVID-19 work.



"Distribution of the publications with different numbers of citations until August 1 of the next calendar year for publications published in 2017, 2018, 2019, and 2020" (Ioannidis et al., 2022)

This citation imbalance created powerful incentives for both researchers and journals. Top medical journals saw unprecedented boosts in impact factor when including COVID-19 papers: for example, *The Lancet*'s impact factor rose 392% (compared to just 35.7% without COVID-19 articles), while other prestigious journals saw similar patterns. The publication rate of COVID-19 manuscripts across all journals saw an increase between 2020 and 2021 with a peak by the second quarter of 2021, with citation rate of COVID-19 manuscripts peaking in the first half of 2020.



Annual impact factor of 6 high-impact medical journals based on (A) manuscripts with and (B) without COVID-19 publications 2019-2021. Changes in impact factor(C-E) with and without COVID-19 publications 2019-2021 (Delardas and Giannos, 2022)

The system created what critics called "an arms race for attention, eyeballs, and citations". As pandemic governance expert Ross Upshur noted, this phenomenon largely represented an intensification of the existing "publish or perish" academic culture rather than a completely new dynamic. Also, because the Impact Factor is weighted so heavily towards highly cited papers, in scenarios like the COVID-19 pandemic, journal IF scores reach extraordinary heights one year, only to crash the next year. This boom and bust cycle can contribute to fuelling further skepticism around Impact Factor in traditional publishing models among researchers.