bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Preprinting the COVID-19 pandemic

Nicholas Fraser[1,#], Liam Brierley[2,#], Gautam Dey[3,4], Jessica K Polka[5], Máté Pálfy[6], Federico Nanni[7] & Jonathon Alexis Coates[8,9,*]

[1] Leibniz Information Centre for Economics, Düsternbrooker Weg 120, 24105 Kiel, Germany

[2] Department of Health Data Science, University of Liverpool, Brownlow Street, Liverpool, L69 3GL, UK

[3] MRC Lab for Molecular Cell Biology, UCL, Gower Street, London WC1E 6BT, UK

[4] Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, Meyerhofstr. 1, 69117 Heidelberg, Germany

[5] ASAPbio, 3739 Balboa St # 1038, San Francisco, CA 94121, USA

[6] The Company of Biologists, Bidder Building, Station Road, Histon, Cambridge CB24 9LF, UK

[7] The Alan Turing Institute, 96 Euston Rd, London NW1 2DB, UK

[8] Hughes Hall College, University of Cambridge, Wollaston Rd, Cambridge, CB1 2EW, UK

[9] William Harvey Research Institute, Charterhouse Square Barts and the London School of Medicine and Dentistry Queen Mary University of London, London, EC1M 6BQ, UK

# These authors contributed equally to this work

* Correspondence: jonathon.coates@qmul.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

The world continues to face an ongoing viral pandemic that presents a serious threat to human health. The virus underlying the COVID-19 disease, SARS-CoV-2, caused over 29 million confirmed cases and 925,000 deaths since January 2020. Although the last pandemic occurred only a decade ago, the way science operates and responds to current events has experienced a paradigm shift in the interim. The scientific community responded rapidly to the COVID-19 pandemic, releasing over 16,000 COVID-19 scientific articles within 4 months of the first confirmed case, of which 6,753 were hosted by preprint servers. Focussing on bioRxiv and medRxiv, two growing preprint servers for biomedical research, we investigated the attributes of COVID-19 preprints, their access and usage rates and characteristics of sharing across online platforms. Our results highlight the unprecedented role of preprint servers in the dissemination of COVID-19 science, and the impact of the pandemic on the scientific communication landscape.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Since January 2020, the world has been gripped by the COVID-19 outbreak, which has escalated to pandemic status, and caused over 29 million cases and 925,000 deaths (3.2 million cases and 220,000 deaths within 4 months) of the first reported case [1,2]. The causative pathogen was rapidly identified as a novel virus within the family *Coronaviridae* and was named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [3]. Although multiple coronaviruses are ubiquitous among humans and cause only mild disease, epidemics of newly emerging coronaviruses were previously observed in SARS in 2002 [4] and Middle East respiratory syndrome (MERS) in 2012 [5]. The unprecedented extent and rate of spread of COVID-19 has created a critical global health emergency and academic communities have raced to respond through research developments.

New research has traditionally been communicated via published journal articles or conference presentations. Traditional scientific publishing involves the submission of manuscripts to an individual journal, which then organises peer review. Authors often conduct additional experiments or analyses to address the reviewers' concerns in one or more revisions. Even after this lengthy process is concluded, almost half of submissions are rejected and require re-submission to a different journal [6]. The median time between the date a preprint is posted and the date on which the first DOI of a journal article is registered is 166 days in the life sciences [7]. Escalating demands made by reviewers and editors are lengthening the publication process still further [8,9].

Preprints are publicly-accessible scientific manuscripts that have not yet been certified by peer review which have been used in other disciplines for over 30 years [10]. In 2013 two new preprint initiatives launched: PeerJ Preprints, from the publisher PeerJ, and bioRxiv, from Cold Spring Harbor Laboratory (CSHL). The latter established partnerships with journals that enabled simultaneous preprint posting at the time of submission [11]. More recently, CSHL, in collaboration with Yale and BMJ, launched medRxiv, a preprint server for the medical sciences [12]. Preprint platforms serving the life sciences have subsequently flourished, with two-thirds of preprints eventually being published in peer-reviewed journals [7].

While funders and institutions explicitly encouraged pre-publication data sharing in the context of the recent Zika and Ebola virus disease outbreaks [13], usage of preprints remained modest through these epidemics [14]. The COVID-19 crisis represents the first time that preprints have been widely used to communicate during an epidemic.

We assessed the role of preprints in the early months of the COVID-19 pandemic, between January 1st and April 30th. We found that preprint servers hosted a large amount of COVID-19 related science, which was being accessed and downloaded in far greater volume than other preprints on the same

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

servers and that these were widely shared across multiple online platforms. Moreover, we determined that COVID-19 preprints are shorter and are reviewed faster than their non-COVID-19 counterparts. Taken together, our data demonstrates the importance of rapidly and openly sharing science in the context of a global pandemic and the essential role of preprints in this endeavour.

Results

COVID-19 preprints were posted early in the pandemic

The COVID-19 pandemic has rapidly spread across the globe, from 3 patients in the city of Wuhan on the 27th December 2019 to over 3.2 million confirmed cases worldwide by the end of April 2020 (Fig. 1A). The World Health Organisation (WHO) declared COVID-19 a pandemic on 11th March [15] as the number of cases grew exponentially throughout the month, despite interventions by governments [16]. The scientific community responded rapidly as COVID-19 emerged as a serious threat, with publications appearing within weeks of the first reported cases (Fig. 1B). By the end of January 2020, 186 scientific articles related to COVID-19 had been published in either a peer-reviewed journal or on a preprint server. When compared to other recent outbreaks of global significance caused by emerging RNA viruses, the response to COVID-19 has been much more rapid; 2,527 COVID-19 related preprints were posted to bioRxiv and medRxiv in the first 4 months of the outbreak alone; in comparison, only 78 Zika virus-related, and 10 Ebola virus-related preprints were posted to bioRxiv and medRxiv during the entire duration of the respective Zika virus epidemic (2015-2016) and Western African Ebola virus epidemic (2014-2016). This surge in COVID-19 preprints is not explained by general increases in preprint server usage; period of epidemic (COVID-19, Ebola or Zika virus) was significantly associated with preprint type (epidemic-related or non-epidemic-related) (Chi-square; $\chi^2 = 1641.6$, df = 2, $p < 0.001$), with the proportion of epidemic-related preprints being greatest for COVID-19 (Supplemental Fig. 1A).

By the end of April more than 16,000 COVID-19 scientific articles had been published. A large proportion of these articles (6,753) were manuscripts hosted on a range of preprint servers (Fig. 1C, data from [17]). Despite being one of the newest preprint servers, medRxiv hosted the largest number of preprints (1,963), whilst other preprint servers (with the exception of SSRN which hosts social sciences and humanities preprints) were each found to host <1,000 preprints (Fig. 1C). Eleven of the 33 preprint servers studied hosted over 100 COVID-19 related preprints each, and the total number of manuscripts is likely an underestimation of the true volume of preprints published as a number of preprint servers that could be expected to host COVID-19 research are not included [17].

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Following a steep increase in the posting of COVID-19 research, traditional publishers adopted new policies to support the ongoing public health emergency response efforts. After multiple public calls from scientists [18], over 30 publishers agreed to make all COVID-19 work freely accessible by the 16th March [19,20]. Shortly after this, publishers (for example eLife [21]) began to alter peer-review policies in an attempt at fast-tracking COVID-19 research. Towards the end of April, OASPA issued an open letter of intent to maximise the efficacy of peer review [22]. The number of open-access COVID-19 journal articles suggests that journals have largely been successful at implementing these new policies (Supplemental Fig. 1B).

Attributes of COVID-19 preprints posted between January and April 2020

To explore the attirbutes of COVID-19 preprints in greater detail, we focused our following investigation on two of the most popular preprint servers in the biomedical sciences: bioRxiv and medRxiv.

Between January and April 2020, 14,812 preprints were deposited in total to bioRxiv and medRxiv, of which the majority (12,285, 82.9%) were non-COVID-19 related preprints (Fig. 2A). While the weekly numbers of non-COVID-19 preprints did not change much during this period, COVID-19 preprint posting increased, peaking at over 250 in early April. When the data was broken down by server, it was evident that whilst posting of COVID-19 preprints to bioRxiv had remained relatively steady, preprints posted to medRxiv increased with time (Supplemental Fig. 2A).

The increase in the rate of preprint posting poses challenges for their timely screening. Only marginally faster screening was detected for COVID-19 preprints than for non-COVID-19 preprints (Fig. 2B) when adjusting for differences between servers (two-way ANOVA, interaction term; $F_{1,14808} = 69.13$, $p < 0.001$). Whilst COVID-19 preprints were screened < 1 day quicker from mean differences observed within both servers (Tukey HSD; both $p < 0.001$), larger differences were observed between servers (Supplemental Fig. 2B), with bioRxiv screening preprints on approximately 2 days quicker than medRxiv for both preprint types (both $p < 0.001$).

The number of authors may give an indication as to the amount of work, resources used, and the extent of collaboration in a paper. While the average number of authors of COVID-19 and non-COVID-19 preprints did not differ, COVID-19 preprints showed slightly more variability in authorship team size (median, 6 [IQR 8] vs 6 [IQR 5]). Single-author preprints were almost three times more common among COVID-19 than non-COVID-19 preprints (Fig. 2C).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Researchers may be shifting their publishing practice in response to the pandemic. Among all identified corresponding authors of preprints during the pandemic, we found a significant association between preprint type and whether this was the author's first bioRxiv or medRxiv preprint (Chi-square, $\chi^2 = 215.2$, df = 1, $p < 0.001$). Among COVID-19 corresponding authors, 83% were posting a preprint for the first time, compared to 68% of non-COVID-19 corresponding authors in the same period. To further understand which authors have been drawn to begin using preprints since the pandemic began, we additionally stratified these groups by country. Corresponding authors based in China had the greatest increase in representation among authors of COVID-19 preprints compared to non-COVID-19 preprints as an expectation (Fig. 2D). Additionally, India had a higher representation among COVID-19 authors specifically using preprints for the first time compared to non-COVID-19 posting patterns. Moreover, we found that most countries posted their first COVID-19 preprint close to the time of their first confirmed COVID-19 case (Supplemental Fig. 2C), with weak positive correlation considering calendar days of both events (Spearman's rank; $\rho = 0.39$, $p = 0.001$). Countries posting a COVID-19 preprint in advance of their first confirmed case were mostly higher-income countries (e.g. USA, UK, New Zealand, Switzerland). COVID-19 preprints were deposited from every inhabited continent, highlighting the global response to the pandemic.

There has been much discussion regarding the appropriateness of researchers switching to COVID-19 research from other fields [33]. To quantify whether this phenomenon was detectable within the preprint literature, we compared the bioRxiv or medRxiv category of each COVID-19 preprint to the most recent previous non-COVID-19 preprint (if any) from the same corresponding author. Most corresponding authors were not drastically changing fields, with category differences generally spanning reasonably related areas (for example, some authors previously posting preprints in evolutionary biology have posted COVID-19 preprints in microbiology) (Supplemental Fig. 2D). This suggests that – at least within the life sciences – principal investigators are utilising their labs' skills and resources in an effective manner in their contributions to COVID-19 research.

bioRxiv and medRxiv allow authors to select from a number of different Creative Commons (https://creativecommons.org/) license types when depositing their work: CC0 (No Rights Reserved), CC-BY (Attribution), CC BY-NC (Attribution, Non-Commercial), CC-BY-ND (Attribution, No-Derivatives), CC-BY-NC-ND (Attribution, Non-Commercial, No-Derivatives). Authors may also select to post their work without a license (i.e. All Rights Reserved). A previous analysis has found that bioRxiv authors tend to post preprints under the more restrictive license types [23], although there appears to be some confusion amongst authors as to the precise implications of each license type [24]. License chosen was significantly associated with preprint type (Fisher's exact, 1000 simulations; $p < 0.001$);

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

authors of COVID-19 preprints were more likely to choose the more restrictive CC-BY-NC-ND or CC-BY-ND than those of non-COVID-19 preprints, and less likely to choose CC-BY and CC (Fig. 2E).

Preprint servers offer authors the opportunity to post new versions of a preprint, to improve upon or correct mistakes in an earlier version. The majority of preprints existed as only a single version for both COVID-19 and non-COVID-19 work with very few preprints existing beyond two versions (Fig. 2F). This may somewhat reflect the relatively short time-span of our analysis period. COVID-19 preprints did not discernibly differ in number of versions compared with non-COVID-19 preprints (median, 1 [IQR 1] vs 1 [IQR 0]).

The speed with which COVID-19 preprints are being posted suggests that researchers have changed the way in which they share results. To investigate this, we compared the word counts of COVID-19 preprints and non-COVID-19 preprints from bioRxiv. We found that COVID-19 preprints are on average 44% shorter in length than non-COVID-19 preprints (median, 3432 [IQR 2597] vs 6143 [IQR 3363]; Mann-Whitney, $p < 0.001$) (Fig. 2G). This supports anecdotal observations that preprints are being used to share more works-in-progress rather than complete stories. We also found that COVID-19 preprints contain fewer references than non-COVID-19 preprints (median, 30.5 [IQR 29] vs 51 [IQR 31]; $p < 0.001$), reflecting the new, emerging COVID-19 field and dearth of prior literature to reference (Fig. 2H).

Critics have previously raised concerns that by forgoing the traditional peer-review process, preprint servers may become flooded by poor-quality research. Nonetheless, several analyses have shown that a large proportion of preprints (~70%) are eventually published in peer-reviewed scientific journals [7]. We assessed differences in publication outcomes for COVID-19 versus non-COVID-19 preprints during our analysis period, which may be partially related to differences in preprint quality. Published status (published/unpublished) was significantly associated with preprint type (Chi-square; $\chi2 = 6.77$, $df = 1$, $p = 0.009$); within our timeframe, 4% of COVID-19 preprints were published by the end of April, compared to the 3% of non-COVID preprints that were published (Fig. 2I). These published COVID-19 preprints were split across many journals, with clinical or multidisciplinary journals surveyed tending to publish the most papers that were previously preprints (Supplemental Fig. 2E). To determine how publishers were prioritising COVID-19 research, we compared the time from preprint posting to publication in a journal. Delay from posting to subsequent publication was significantly accelerated for COVID-19 preprints by a mean difference of 26.2 days compared to non-COVID-19 preprints posted in the same time period (mean, 22.5 days [SD 15.7] vs 48.7 days [SD 25.6]; two-way ANOVA; $F_{1,289} = 69.8$, $p < 0.001$). This did not appear driven by any temporal changes in publishing practices, as publication times of non-COVID-19 preprints were similar to expectation of our control timeframe of

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

September - January (Fig. 2J). COVID-19 preprints also appeared to have significantly accelerated publishing regardless of publisher (two-way ANOVA, interaction term; $F_{6,283} = 0.41$, $p = 0.876$) (Supplemental Fig. 2F). However, data aggregated across several publishers revealed that on average, non-COVID-19 manuscripts had a 10.6% higher acceptance rate than COVID-19 manuscripts, regardless of preprint availability (Supplemental Fig. 2G).

Extensive access of preprint servers for COVID-19 research

Throughout our time window, COVID-19 preprints received abstract views at a rate over 15 times that of non-COVID-19 preprints (Fig. 3A) (time-adjusted negative binomial regression; rate ratio = 15.6, $z = 143.8$, $p < 0.001$) and downloads at a rate of almost 30 times (Fig. 3B) (rate ratio = 28.9, $z = 155.1$, $p < 0.001$). Abstract views and downloads also appeared to slightly reduce over time for all preprints, with each additional calendar week in posting date resulting in a 6.3% reduction in rate of views (rate ratio = 0.937, $z = -44.56$, $p < 0.001$) and an 8.1% reduction in rate of downloads (rate ratio = 0.919, $z = -51.07$, $p < 0.001$), i.e., most preprints received their heaviest usage near to time of posting, but slowly continued to accumulate usage over time, the highest rates of usage being observed for COVID-19 preprints posted during the week commencing 20th January.

To confirm that usage of COVID-19 and non-COVID-19 preprints was not an artefact of differing preprint server reliance during the pandemic, we compared usage to September 2019 – April 2020, as a non-pandemic control period. We observed a slight decrease in abstract views (Supplemental Fig. 3A) and pdf downloads (Supplemental Fig. 3B) in March 2020, but otherwise, the usage data did not differ from that prior to the pandemic.

Secondly, we investigated usage across additional preprint servers (data kindly provided by each of the server operators). We found that COVID-19 preprints were consistently downloaded more than non-COVID-19 preprints during our timeframe, regardless of which preprint server hosted the manuscript (Supplemental Fig. 3C), though the gap in downloads varied between server (two-way ANOVA, interaction term; $F_{4,276544} = 586.9$, $p < 0.001$). Server usage differences were more pronounced for COVID-19 preprints; multiple post-hoc comparisons confirmed that bioRxiv and medRxiv received significantly higher usage per COVID-19 preprint than all other servers for which data was available (Tukey HSD; all p values < 0.001). However, for non COVID-19 preprints, the only observed pairwise differences between servers indicated greater bioRxiv usage than SSRN or Research Square (Tukey HSD; all p values < 0.001). This suggests specific attention has been given disproportionately to bioRxiv and medRxiv as repositories for COVID-19 research.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

COVID-19 preprints were shared more widely than non-COVID-19 preprints

Based on citation data from Dimensions, we found that COVID-19 preprints are cited more often than non-COVID-19 preprints (time-adjusted negative binomial regression; rate ratio = 71.1, $z = 49.2$, $p < 0.001$) (Fig. 4A), although it should be noted that only a minority of preprints received at least one citation in both groups (30.6% vs 5.5%). The highest cited preprint had 127 citations, with the 10th most cited COVID-19 preprint receiving 48 citations (Table 1); many of the highest cited preprints focussing on the viral cell receptor, angiotensin converting enzyme 2 (ACE2) or the epidemiology of COVID-19.

We also investigated sharing of preprints on Twitter to assess the exposure of wider public audiences to preprints, using data from Altmetric. COVID-19 preprints were tweeted at a greater rate than non-COVID-19 preprints (rate ratio = 14.8, $z = 91.55$, $p < 0.001$) (Fig. 4B). The most tweeted non-COVID-19 preprint received 1,323 tweets, whereas 8 of the top 10 tweeted COVID-19 preprints were tweeted over 10,000 times each (Table 2). Many of the top 10 tweeted COVID-19 preprints were related to transmission, re-infection or seroprevalence and association with the BCG vaccine. The most tweeted COVID-19 preprint (29,984 tweets) was a study investigating antibody seroprevalence in California [25], whilst the second most tweeted COVID-19 preprint was a widely criticised (and later withdrawn) study linking the SARS-CoV-2 spike protein to HIV-1 glycoproteins [26].

To better understand the main discussion topics associated with the top-10 most tweeted preprints, we analysed the hashtags used in original tweets (i.e. excluding retweets) mentioning those preprints (Supplemental Fig. 4A). After removing generic or overused hashtags directly referring to the virus (e.g. “#coronavirus”, “#COVID-19”), we found that the most dominant hashtag among tweets referencing preprints was “#hydroxychloroquine”, a major controversial topic associated with two of the top ten most tweeted preprints. Other prominent hashtags contained a mixture of direct, neutral references to the disease outbreak such as “#coronavirusoutbreak” and “#Wuhan”, and some more politicised terms, such as “#fakenews” and “#covidisalie”, associated with conspiracy theories.

As well as featuring heavily on social media, COVID-19 research has also pervaded print and online news media. COVID-19 preprints were used in news articles at a rate over two hundred times that of non-COVID-19 preprints (rate ratio = 220.4, $z = 39.27$, $p < 0.001$), although as with citations, only a minority were mentioned in news articles at all (26.9% vs 6.7%) (Fig. 4C). The top 10 non-COVID-19 preprints were reported in less than 100 news articles in total, whereas the top COVID-19 preprints were reported in over 300 news articles (Table 3). Similarly, COVID-19 preprints were also used in

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

blogs at a significantly greater rate than non-COVID-19 preprints (rate ratio = 9.48, $z = 29.2$, $p < 0.001$) (Fig. 4D; Table 4). We noted that several of the most widely-disseminated non-COVID-19 preprints featured topics relevant to infectious disease research, e.g. human respiratory physiology and personal protective equipment (Tables 2 and 3).

Independent COVID-19 review projects have arisen to publicly review COVID-19 preprints [34]. To investigate engagement with preprints directly on the bioRxiv and medRxiv platforms, we quantified the number of comments for preprints posted between January and April. We found that non-COVID-19 preprints were rarely commented upon when compared to COVID-19 preprints (time-adjusted negative binomial regression; rate ratio = 27.9, $z = 32.0$, $p < 0.001$) (Fig. 2E); the most commented non-COVID-19 preprint received only 15 comments, whereas the most commented COVID-19 preprint had over 500 comments on the 30th April (Table 5). One preprint, which had 127 comments was retracted within 3 days of being posted following intense public scrutiny [35]. Collectively these data suggest that the most discussed or controversial COVID-19 preprints are being rapidly and publicly scrutinised, with commenting systems being used for direct feedback and discussion of preprints.

Among a set of 66 COVID-19 policy documents (which were manually retrieved from the European Centre for Disease Prevention and Control (ECDC), National Academy of Medicine (NAM), United Kingdom Parliamentary Office of Science and Technology (UK POST), United States House Select Subcommittee on the Coronavirus Crisis (US HSSCC), and World Health Organisation Scientific Briefs (WHO SB)), 26 documents cited preprints (including servers beyond bioRxiv and medRxiv, e.g. SSRN, Research Square, arXiv). However, these citations occurred at a relatively low rate, typically constituting less than 20% of the total citations in these 26 documents (Fig. 2F). Fifty-eight individual COVID-19 preprints from bioRxiv or medRxiv were cited in examined policy documents, of which 17 were cited more than once and 4 were cited more than twice. Most preprint citations occurred in documents from the ECDC, UK POST and WHO SB with no preprints cited in analysed documents from the US HSSCC. In comparison, only two instances of citations to preprints were observed among 26 manually collected non-COVID-19 policy documents from the same sources.

To understand how different usage indicators may represent the sharing behaviour of different user groups, we calculated the correlation between the usage indicators presented above (citations, tweets, news articles, comments). For COVID-19 preprints, we found weak correlation between the numbers of citations and Twitter shares (Spearman's $\rho = 0.37$, $p < 0.001$), and the numbers of citations and news articles (Spearman's $\rho = 0.41$, $p < 0.001$) (Fig. 4G), suggesting that the preprints cited mostly

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

within the scientific literature differed to those that were mostly shared by the wider public on other online platforms. There was a stronger correlation between COVID-19 preprints that were most blogged and those receiving the most attention in the news (Spearman's $\rho = 0.58$, $p < 0.001$). Moreover, there was a strong correlation between COVID-19 preprints that were most tweeted and those receiving the most attention in the news (Spearman's $\rho = 0.53$, $p < 0.001$), suggesting similarity between preprints shared on social media and in news media (Fig. 4G). There was a weak correlation between the 10 most tweeted COVID-19 preprints and the 10 most commented upon (Spearman's $\rho = 0.41$, $p < 0.001$). Taking the top ten COVID-19 preprints by each indicator, there was substantial overlap between all indicators except citations (Supplemental Fig. 4B). We observed much weaker correlation between all indicators for non-COVID-19 preprints (Fig. 4H).

Our data reveals that COVID-19 preprints received a significant amount of attention from scientists, news organisations, the general public and policy making bodies, representing a departure for how preprints are normally shared (considering observed patterns for non-COVID-19 preprints).

Table 1. Top 10 cited COVID-19 preprints

Table 2. Top 10 tweeted COVID-19 preprints

Table 3. Top 10 COVID-19 preprints covered by news organisations

Table 4. Top 10 most commented COVID-19 preprints

Table 5. Top 10 most blogged COVID-19 preprints

Discussion

Usage of preprint servers within the biological sciences has been rising since the inception of bioRxiv and other platforms [10,27]. The urgent health threat of a global pandemic has catapulted the use of preprint servers as a means of quickly disseminating scientific findings into the public sphere, encouraged by funding bodies requiring COVID-19 research to be open access [18,20]. Our results show that preprints have been widely adopted for the dissemination and communication of COVID-19 research, and in turn, the pandemic has greatly impacted the preprint and science publishing landscape.

Changing attitudes and acceptance within the life sciences to preprint servers may be one reason why COVID-19 research is being shared to readily as preprints compared to past epidemics. In addition, the

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

need to rapidly communicate findings prior to a lengthy review process might be responsible for this observation (Fig. 3). A recent study involving qualitative interviews of multiple research stakeholders found "early and rapid dissemination" to be amongst the most often cited benefits of preprints [28]. These findings were echoed in a survey of ~4200 bioRxiv users [10], and are underscored by the 6 month median lag between posting of a preprint and subsequent journal publication [7,27]. Such timelines for disseminating findings are clearly incompatible with the lightning-quick progression of a pandemic. An analysis of publication timelines for 14 medical journals has shown that some publishers have taken steps to accelerate their publishing processes for COVID-19 research, reducing the time for the peer-review stage (submission to acceptance) on average by 45 days, and the editing stage (acceptance to publication) by 14 days [29], yet this still falls some way short of the ~1-3 day screening time for bioRxiv and medRxiv preprints (Fig. 2B). Further studies on understanding the motivations behind posting preprints, for example through quantitative and qualitative author surveys may help funders and other stakeholders that support the usage of preprints to address some of the social barriers for their uptake [30].

bioRxiv and medRxiv included a banner to explain that preprints should not be regarded as conclusive and not reported on in the news media as established information [31]. Despite the warning message, COVID-19 preprints have received unprecedented coverage on online media platforms (Fig. 4). Twitter has been a particularly notable outlet for communication of preprints, a finding echoed by a recent study on the spread of the wider (i.e. not limited to preprints) COVID-19 research field on Twitter, which found that COVID-19 research was being widely disseminated and driven largely by academic Twitter users [32]. Nonetheless, the relatively weak correlation found between citations and other indicators of online sharing (Fig 4G) suggests that the interests of scientists versus the broader public largely differ: of the articles in the top-10 most shared on twitter, in news articles or on blogs, only one is ranked amongst the top-10 most cited articles (Supplemental Fig. 4B). Hashtags associated with individual, highly tweeted preprints reveal some emergent themes that suggest communication of certain preprints can also extend well beyond scientific audiences (Supplemental Fig. 4A). These range from good public health practice ("#washyourhands") to right-wing philosophies and conspiracy theories, ("#fakenews" and "#endthelockdown"). This type of misinformation is common to new diseases [33] and social media platforms have recently released a statement outlining their plans to combat this issue [34]. An even greater adoption of open science principles has recently been suggested as one method to counter such misuse of preprints and peer-reviewed articles [35], though for now, this remains an increasingly important discourse.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The fact that news outlets are reporting extensively on COVID-19 preprints (Fig. 4C and 4G) represents a marked change in journalistic practice: pre-pandemic, bioRxiv preprints received very little coverage in comparison to journal articles [27]. This cultural shift provides an unprecedented opportunity to bridge the scientific and media communities to create a consensus on the reporting of preprints [36]. Another marked change was observed in the use of preprints in policy documents (Fig. 4F). Preprints were remarkably absent in non-COVID-19 policy documents yet present, albeit at relatively low levels, in COVID-19 policy documents. In a larger dataset, two of the top 10 journals which are being cited in policy documents were found to be preprint servers (medRxiv and SSRN in 5th and 8th position respectively) [37]. This suggests that preprints are being used to directly influence policy-makers and decision making. We only investigated a limited set of policy documents, largely restricted to Europe and the US and whether this extends more globally remains to be explored. In the near future, we aim to examine the use of preprints in policy in more detail to address these questions.

As most COVID-19-preprints were not yet published, concerns regarding quality will persist [38]. Despite increased publicity for established preprint-review services (such as PREreview [22,39]), there has been limited use of these platforms [40]. However, independent preprint-review projects have arisen whereby reviews are posted in the comments section of preprint servers or hosted on independent websites [41,42]. These more formal projects partly account for the increased commenting on COVID-19 preprints (Fig. 4). Moreover, prominent scientists are using social media platforms such as Twitter to publicly share concerns with poor quality COVID-19 preprints or to amplify high-quality preprints [43]. The use of Twitter to “peer-review” preprints provides additional, public, scrutiny on manuscripts that can complement the less opaque and slower traditional peer-review process. Although these new review platforms partially combat poor-quality preprints, it is clear that there is a dire need to better understand the general quality and trustworthiness of preprints compared to peer-review articles. We found comparative levels of preprints had been published within our short timeframe (Fig. 2) and that acceptance rates at several journals was only slightly reduced for COVID-19 research compared to non-COVID-19 articles (Supplemental Fig. 2) suggesting that, generally, preprints were relatively of good quality. Furthermore, recent studies have suggested that the quality of reporting in preprints differs little from their later peer-reviewed articles [44] and we ourselves are currently undertaking a more detailed analysis (see version 1 of our preprint for an initial analysis of published COVID preprints [45]). However, the problem of poor-quality science is not unique to preprints and ultimately, a multi-pronged approach is required to solve some of these issues. For example, scientists must engage more responsibly with journalists and the public, in addition to upholding high standards when sharing research. More significant consequences for academic misconduct and the swift removal of problematic articles will be essential in aiding this.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Moreover, the politicisation of science has become a polarising issue and must be prevented at all costs. Thirdly, transparency within the scientific process is essential in improving the understanding of its internal dynamics and providing accountability.

Our data demonstrates the indispensable role that preprints, and preprint servers, are playing during a global pandemic. By communicating science through preprints, we are sharing at a faster rate than allowed by the current journal infrastructure. Furthermore, we provide evidence for important future discussions around scientific publishing.

Methods

Preprint Metadata for bioRxiv and medRxiv

We retrieved basic preprint metadata (DOIs, titles, abstracts, author names, corresponding author name and institution, dates, versions, licenses, categories and published article links) for bioRxiv and medRxiv preprints via the bioRxiv Application Programming Interface (API; https://api.biorxiv.org). The API accepts a 'server' parameter to enable retrieval of records for both bioRxiv and medRxiv. We initially collected metadata for all preprints posted from the time of the server's launch, corresponding to November 2013 for bioRxiv and June 2019 for medRxiv, until the end of our analysis period on 30th April 2020 (N = 84,524). All data were collected on 1st May 2020. Note that where multiple preprint versions existed, we included only the earliest version and recorded the total number of following revisions. Preprints were classified as "COVID-19 preprints" or "non-COVID-19 preprints" on the basis of the following terms contained within their titles or abstracts (case-insensitive): "coronavirus", "covid-19", "sars-cov", "ncov-2019", "2019-ncov", "hcov-19", "sars-2". For comparison of preprint behaviour between the COVID-19 outbreak and previous viral epidemics, namely Western Africa Ebola virus and Zika virus (Supplemental Fig. 1), the same procedure was applied using the keywords "ebola" or "zebov", and "zika" or "zikv", respectively.

For a subset of preprints posted between 1st September 2019 and 30th April 2020 (N = 25,883), we enhanced the basic preprint metadata with data from a number of other sources, as outlined below. Note that this time period was chosen to encapsulate our 4-month analysis period from 1st January to 30th April 2020 (N = 14,812), as well as the preceding 4-month period from September 1st to December 31st 2019 (N = 11,071), to use for comparison purposes. Of the preprints contained in the later 4-month analysis period, 2,527 (17.1%) contained COVID-19 related keywords in their titles or abstracts.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

For all preprints contained in the subset, disambiguated author affiliation and country data for corresponding authors were retrieved by querying raw affiliation strings against the Research Organisation Registry (ROR) API (https://github.com/ror-community/ror-api). The API provides a service for matching affiliation strings against institutions contained in the registry, on the basis of multiple matching types (named “phrase”, “common terms”, “fuzzy”, “heuristics”, and “acronyms”). The service returns a list of potential matched institutions and their country, as well as the matching type used, a confidence score with values between 0 and 1, and a binary “chosen” indicator relating to the most confidently matched institution. A small number (~500) of raw affiliation strings returned from the bioRxiv API were truncated at 160 characters; for these records we conducted web-scraping using the rvest package for R [46] to retrieve the full affiliation strings of corresponding authors from the bioRxiv public webpages, prior to matching. For the purposes of our study, we aimed for higher precision than recall, and thus only included matched institutions where the API returned a confidence score of 1. A manual check of a sample of returned results also suggested higher precision for results returned using the “phrase” matching type, and thus we only retained results using this matching type. In a final step, we applied manual corrections to the country information for a small subset of records where false positives would be most likely to influence our results by a) iteratively examining the chronologically first preprint associated with each country following affiliation matching and applying manual rules to correct mismatched institutions until no further errors were detected (n = 8 institutions); and b) examining the top 50 most common raw affiliation strings and applying manual rules to correct any mismatched or unmatched institutions (n = 2 institutions). In total, we matched 19,002 preprints to a country (73.2%); for COVID-19 preprints alone, 1716 preprints (67.9%) were matched to a country. Note that a similar, albeit more sophisticated method of matching bioRxiv affiliation information with the ROR API service was recently documented by Abdill et al. [47].

Word counts and reference counts for each preprint were also added to the basic preprint metadata via scraping of the bioRxiv public webpages (medRxiv currently does not display full HTML texts, and so calculating word and reference counts was limited to bioRxiv preprints). Web scraping was conducted using the rvest package for R [46]. Word counts refer to words contained only in the main body text, after removing the abstract, figure captions, table captions, acknowledgements and references. In a small number of cases, word counts could not be retrieved because no full-text existed; this occurs as we targeted only the first version of a preprint, but in cases where a second version was uploaded very shortly (i.e. within a few days) after the first version, the full-text article was generated only for the second version. Word and reference counts were retrieved for 21,975 of 22,156 bioRxiv preprints (99.1%); for COVID-19 preprints alone, word and reference counts were

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

retrieved for 553 of 564 preprints (98.0 %). Word counts ranged from 583 to 39,953 words, whilst reference counts ranged from 1 to 487 references.

Our basic preprint metadata retrieved from the bioRxiv API also contained DOI links to published versions (i.e. a peer-reviewed journal article) of preprints, where available. In total, 2710 records in our preprint subset (10.5%) contained links to published articles, although of COVID-19 preprints only 101 preprints contained such links (4.0%). It should be noted that COVID-19 articles are heavily weighted towards the most recent months of the dataset and have thus had less time to progress through the journal publication process. Links to published articles are likely an underestimate of the total proportion of articles that have been subsequently published in journals – both as a result of the delay between articles being published in a journal and being detected by bioRxiv, and bioRxiv missing some links to published articles when e.g. titles change significantly between the preprint and published version [27]. Published article metadata (titles, abstracts, publication dates, journal and publisher name) were retrieved by querying each DOI against the Crossref API (https://api.crossref.org), using the rcrossref package for R [48]. We also retrieved data regarding the open access status of each article by querying each DOI against the Unpaywall API, via the roadoi package for R [49].

Usage, Altmetrics and Citation Data

For investigating the rates at which preprints are used, shared and cited, we collected detailed usage, altmetrics and citation data for all bioRxiv and medRxiv preprints posted between 1st September 2019 to 30th April 2020 (i.e. for every preprint where we collected detailed metadata, as described in the previous section). Collection of all usage, altmetrics and citation data were conducted on 1st May 2020.

Usage data (abstract views and pdf downloads) were scraped from the bioRxiv and medRxiv public webpages, using the rvest package for R (Wickham, 2019). bioRxiv and medRxiv webpages display abstract views and pdf downloads on a calendar month basis; for subsequent analysis (e.g Figure 4), these were summed to generate total abstract views and downloads since the time of preprint posting. In total, usage data were recorded for 25,865 preprints (99.9%) – a small number were not recorded, possibly due to server issues during the web scraping process. Note that bioRxiv webpages also display counts of full-text views, although we did not include these data in our final analysis. This was partially to ensure consistency with medRxiv, which currently does not provide display full HTML texts, and partially due to ambiguities in the timeline of full-text publishing – the full text of a preprint is added several days after the preprint is first available, but the exact delay appears to vary from preprint to preprint. We also compared rates of PDF downloads for bioRxiv and medRxiv preprints with a number of other preprint servers (Preprints.org, SSRN, and Research Square) (Supplemental

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig. 3C) - these data were provided directly by representatives of each of the respective preprint servers.

Counts of multiple altmetric indicators (mentions in tweets, blogs, and news articles) were retrieved via Altmetric (https://www.altmetric.com), a service that monitors and aggregates mentions to scientific articles on various online platforms. Altmetric provide a free API (https://api.altmetric.com) against which we queried each preprint DOI in our analysis set. Importantly, Altmetric only contains records where an article has been mentioned in at least one of the sources tracked, thus, if our query returned an invalid response we recorded counts for all indicators as zero. Coverage of each indicator (i.e. the proportion of preprints receiving at least a single mention in a particular source) for preprints were 99.1%, 9.6%, and 3.5% for mentions in tweets, blogs and news articles respectively. The high coverage on Twitter is likely driven, at least in part, by automated tweeting of preprints by the official bioRxiv and medRxiv twitter accounts. For COVID-19 preprints, coverage was found to be 100.0%, 16.6% and 26.9% for mentions in tweets, blogs and news articles respectively.

To quantitatively capture how high-usage preprints were being received by Twitter users, we retrieved all tweets linking to the top ten most-tweeted preprints. Tweet IDs were retrieved via the Altmetric API service, and then queried against the Twitter API using the rtweet package [50] for R, to retrieve full tweet content.

Citations counts for each preprint were retrieved from the scholarly indexing database Dimensions (https://dimensions.ai). An advantage of using Dimensions in comparison to more traditional citation databases (e.g. Scopus, Web of Science) is that Dimensions also includes preprints from several sources within their database (including from bioRxiv and medRxiv), as well as their respective citation counts. When a preprint was not found, we recorded its citation counts as zero. Of all preprints, 3707 (14.3%) recorded at least a single citation in Dimensions. For COVID-19 preprints, 774 preprints (30.6%) recorded at least a single citation.

Comments

BioRxiv and medRxiv html pages feature a Disqus (https://disqus.com) comment platform to allow readers to post text comments. Comment counts for each bioRxiv and medRxiv preprint were retrieved via the Disqus API service (https://disqus.com/api/docs/). Where multiple preprint versions existed, comments were aggregated over all versions. As with preprint perceptions among public audiences on Twitter, we then examined perceptions among academic audiences by examining comment sentiment. Text content of comments for COVID-19 preprints were provided directly by the bioRxiv development team.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Screening time for bioRxiv and medRxiv

To calculate screening time, we followed the method outlined by Steve Royle [51]. In short, we calculate the screening time as the difference in days between the preprint posting date, and the date stamp of submission approval contained within bioRxiv and medRxiv DOIs (only available for preprints posted after December 11th 2019). bioRxiv and medRxiv preprints were filtered to preprints posted between January 1st – April 30th 2020, accounting for the first version of a posted preprint.

Policy documents

To describe the level of reliance upon preprints in policy documents, a set of policy documents were manually collected from the following institutional sources: the European Centre for Disease Prevention and Control (including rapid reviews and technical reports), National Academy of Medicine, UK Parliamentary Office of Science and Technology, US House Select Subcommittee on the Coronavirus Crisis documents, and the WHO (n = 66 COVID-19 related policies, n = 26 non-COVID-19 related policies). COVID-19 policy documents were selected from 1st January 2020 – 17th June 2020. Due to the limited number of non-COVID-19 policy documents from the same time period, these documents were selected dating back to September 2018. Reference lists of each policy document were then text-mined and manually verified to calculate the proportion of references that were preprints.

Statistical analyses

Preprint counts were compared across categories (e.g., COVID-19 or non-COVID-19) using Chi-square tests or, in cases where any expected values were < 5, with Fisher's exact tests using Monte Carlo simulation. Quantitative preprint metrics (e.g. word count, comment count) were compared across categories using Mann-Whitney tests and correlated with other quantitative metrics using Spearman's rank tests for univariate comparisons.

For time-variant metrics (e.g. views, downloads, which may be expected to vary with length of preprint availability), we analysed the difference between COVID-19 and non-COVID-19 preprints using generalised linear regression models with calendar days since Jan 1st 2020 as an additional covariate and negative binomially-distributed errors. This allowed estimates of time-adjusted rate ratios comparing COVID-19 and non-COVID-19 preprint metrics. Negative binomial regressions were constructed using the function 'glm.nb' in R package MASS [52]. For multivariate categorical comparisons of preprint metrics (e.g. screening time between preprint type and preprint server), we constructed two-way factorial ANOVAs, testing for interactions between both category variables in all cases. Pairwise post-hoc comparisons of interest were tested using Tukey's honest significant difference (HSD) while correcting for multiple testing, using function 'glht' in R package multcomp [53].

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Parameters and limitations of this study

We acknowledge a number of limitations in our study. Firstly, to assign a preprint as COVID-19 or not, we used keyword matching to titles/abstracts on the preprint version at the time of our data extraction. This means we may have captured some early preprints, posted before the pandemic that had been subtly revised to include a keyword relating to COVID-19. Our data collection period was a tightly defined window (January-April 2020) which may impact upon the altmetric and usage data we collected as those preprints posted at the end of April would have had less time to accrue these metrics.

Acknowledgements

The authors would like to thank Ted Roeder, John Inglis and Richard Sever from bioRxiv and medRxiv for providing information relating to comments on COVID-19 preprints. We would also like to thank Martyn Rittman (preprints.org), Shirley Decker-Lucke (SSRN) and Michele Avissar-Whiting (Research Square) for kindly providing usage data. Further thanks to Helena Brown and Sarah Bunn for conversations regarding media usage and government policy. NF acknowledges funding from the German Federal Ministry for Education and Research, grant numbers 01PU17005B (OASE) and 01PU17011D (QuaMedFo). LB acknowledges funding from a Medical Research Council Skills Development Fellowship award, grant number MR/T027355/1

Author contributions

Conceptualisation, N.F., L.B., G.D., J.K.P., M.P., J.A.C.; Methodology, N.F., L.B., J.A.C.; Software, N.F., L.B.; Validation, N.F., L.B., J.A.C.; Formal analysis, N.F., L.B., J.A.C.; Investigation, N.F., L.B., G.D., J.K.P., M.P., J.A.C.; Resources, J.K.P. and J.A.C.; Data curation, N.F., L.B., J.A.C.; Writing – original draft, N.F., L.B., G.D., J.K.P., M.P., J.A.C.; Writing – Review & editing, N.F., L.B., G.D., J.K.P., M.P., F.N., J.A.C.; Visualisation, N.F., L.B., J.A.C.; Supervision, J.A.C.; Project administration, J.A.C.

Data availability

All data and code used in this study are available on github (https://github.com/preprinting-a-pandemic/pandemic_preprints), with the exception of data provided by preprint servers, publishers and raw tweet data. Data not publicly available may be shared following permission from the relevant provider and upon reasonable request to the corresponding author.

Declaration of interests

JP is the executive director of ASAPbio, a non-profit organization promoting the productive use of preprints in the life sciences. GD is a bioRxiv Affiliate, part of a volunteer group of scientists that screen preprints deposited on the bioRxiv server. MP is the community manager for preLights, a non-profit

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

preprint highlighting service. GD and JAC are contributors to preLights and ASAPbio Fellows. The authors declare no other competing interests.

References

1. WHO. COVID-19 situation report 19. 2 Aug 2020 [cited 13 May 2020]. Available: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200501-covid-19-sitrep.pdf

2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020;382: 727–733. doi:10.1056/NEJMoa2001017

3. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020;5: 536–544. doi:10.1038/s41564-020-0695-z

4. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. In: http://dx.doi.org/10.1056/NEJMoa030781 [Internet]. Massachusetts Medical Society; 7 Oct 2009 [cited 13 May 2020]. doi:10.1056/NEJMoa030781

5. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. N Engl J Med. 2012;367: 1814–1820. doi:10.1056/NEJMoa1211721

6. Wallach JD, Egilman AC, Gopal AD, Swami N, Krumholz HM, Ross JS. Biomedical journal speed and efficiency: a cross-sectional pilot survey of author experiences. Res Integr Peer Rev. 2018;3: 1. doi:10.1186/s41073-017-0045-8

7. Abdill RJ, Blekhman R. Tracking the popularity and outcomes of all bioRxiv preprints. Pewsey E, Rodgers P, Greene CS, editors. eLife. 2019;8: e45133. doi:10.7554/eLife.45133

8. Powell K. Does it take too long to publish research? Nat News. 2016;530: 148. doi:10.1038/530148a

9. Vale RD. Accelerating scientific publication in biology. Proc Natl Acad Sci. 2015;112: 13439–13446. doi:10.1073/pnas.1511912112

10. Sever R, Roeder T, Hindle S, Sussman L, Black K-J, Argentine J, et al. bioRxiv: the preprint server for biology. bioRxiv. 2019; 833400. doi:10.1101/833400

11. Kaiser J, 2014, Am 12:00. BioRxiv at 1 year: A promising start. In: Science | AAAS [Internet]. 11 Nov 2014 [cited 13 May 2020]. Available: https://www.sciencemag.org/news/2014/11/biorxiv-1-year-promising-start

12. Rawlinson C, Bloom T. New preprint server for medical research. BMJ. 2019;365. doi:10.1136/bmj.l2301

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

13. Wellcome Trust. Sharing data during Zika and other global health emergencies | Wellcome. In: Wellcome.ac.uk [Internet]. 2 Oct 2016 [cited 13 May 2020]. Available: https://wellcome.ac.uk/news/sharing-data-during-zika-and-other-global-health-emergencies

14. Johansson MA, Reich NG, Meyers LA, Lipsitch M. Preprints: An underutilized mechanism to accelerate outbreak science. PLOS Med. 2018;15: e1002549. doi:10.1371/journal.pmed.1002549

15. WHO. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 11 Mar 2020 [cited 13 May 2020]. Available: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

16. Cheng C, Barcelo J, Hartnett A, Kubinec R, Messerschmidt L. CoronaNet: A Dyadic Dataset of Government Responses to the COVID-19 Pandemic. 2020 [cited 27 Apr 2020]. doi:10.31235/osf.io/dkvxy

17. Fraser N, Kramer B. covid19_preprints. 2020. doi:10.6084/m9.figshare.12033672.v14

18. Wellcome Trust. Coronavirus (COVID-19): sharing research data | Wellcome. 31 Jan 2020 [cited 21 May 2020]. Available: https://wellcome.ac.uk/coronavirus-covid-19/open-data

19. Nature Editorial. Calling all coronavirus researchers: keep sharing, stay open. Nature. 2020;578: 7–7. doi:10.1038/d41586-020-00307-x

20. Wellcome Trust. Publishers make coronavirus (COVID-19) content freely available and reusable | Wellcome. 16 Mar 2020 [cited 21 May 2020]. Available: https://wellcome.ac.uk/press-release/publishers-make-coronavirus-covid-19-content-freely-available-and-reusable

21. Eisen MB, Akhmanova A, Behrens TE, Weigel D. Publishing in the time of COVID-19. eLife. 2020;9: e57162. doi:10.7554/eLife.57162

22. OASPA. COVID-19 Publishers Open Letter of Intent - Rapid Review. In: OASPA [Internet]. 27 May 2020 [cited 13 May 2020]. Available: https://oaspa.org/covid-19-publishers-open-letter-of-intent-rapid-review/

23. Himmelstein D. The licensing of bioRxiv preprints. Satoshi Village. 2016 [cited 19 May 2020]. Available: https://blog.dhimmel.com/biorxiv-licenses/

24. ASAPbio. asapbio/licensing. ASAPbio; 2018. Available: https://github.com/asapbio/licensing

25. Bendavid E, Mulaney B, Sood N, Shah S, Ling E, Bromley-Dulfano R, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. medRxiv. 2020; 2020.04.14.20062463. doi:10.1101/2020.04.14.20062463

26. Pradhan P, Pandey AK, Mishra A, Gupta P, Tripathi PK, Menon MB, et al. Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag. bioRxiv. 2020; 2020.01.30.927871. doi:10.1101/2020.01.30.927871

27. Fraser N, Momeni F, Mayr P, Peters I. The relationship between bioRxiv preprints, citations and altmetrics. Quant Sci Stud. 2020; 1–21. doi:10.1162/qss_a_00043

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

28. Chiarelli A, Johnson R, Pinfield S, Richens E. Preprints and Scholarly Communication: An Exploratory Qualitative Study of Adoption, Practices, Drivers and Barriers. F1000Research. 2019;8: 971. doi:10.12688/f1000research.19619.2

29. Horbach SPJM. Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19. Quant Sci Stud. 2020;1: 1056–1067. doi:10.1162/qss_a_00076

30. Penfold NC, Polka JK. Technical and social issues influencing the adoption of preprints in the life sciences. PLOS Genet. 2020;16: e1008565. doi:10.1371/journal.pgen.1008565

31. Inglis J. We've just put an additional, cautionary note about the use of preprints on every @biorxivpreprint https://t.co/08eSXL4dDi. In: Twitter [Internet]. 2 Jan 2020 [cited 22 May 2020]. Available: https://twitter.com/johnringlis/status/1223598414493077505

32. Fang Z, Costas R. Tracking the Twitter attention around the research efforts on the COVID-19 pandemic. ArXiv200605783 Cs. 2020 [cited 16 Sep 2020]. Available: http://arxiv.org/abs/2006.05783

33. Mian A, Khan S. Coronavirus: the spread of misinformation. BMC Med. 2020;18. doi:10.1186/s12916-020-01556-3

34. Lally C, Christie L. COVID-19 misinformation. UK Parliam POST. 2020 [cited 21 May 2020]. Available: https://post.parliament.uk/analysis/covid-19-misinformation/, https://post.parliament.uk/analysis/covid-19-misinformation/

35. Besançon L, Peiffer-Smadja N, Segalas C, Jiang H, Masuzzo P, Smout C, et al. Open Science Saves Lives: Lessons from the COVID-19 Pandemic. bioRxiv. 2020; 2020.08.13.249847. doi:10.1101/2020.08.13.249847

36. Sheldon T. Preprints could promote confusion and distortion. Nature. 2018;559: 445–446.

37. Adie E. COVID-19-policy dataset. 2020. doi:10.6084/m9.figshare.12055860.v2

38. Bagdasarian N, Cross GB, Fisher D. Rapid publications risk the integrity of science in the era of COVID-19. BMC Med. 2020;18: 192. doi:10.1186/s12916-020-01650-6

39. Johansson MA, Saderi D. Open peer-review platform for COVID-19 preprints. Nature. 2020;579: 29–29. doi:10.1038/d41586-020-00613-4

40. Brierley L. The role of research preprints in the academic response to the COVID-19 epidemic. 2020. doi:10.22541/au.158516578.89167184

41. Vabret N, Samstein R, Fernandez N, Merad M. Advancing scientific knowledge in times of pandemics. Nat Rev Immunol. 2020; 1–1. doi:10.1038/s41577-020-0319-0

42. MIT Press. The MIT Press and UC Berkeley launch Rapid Reviews: COVID-19. In: MIT News | Massachusetts Institute of Technology [Internet]. 29 Jun 2020 [cited 13 Sep 2020]. Available: https://news.mit.edu/2020/mit-press-and-uc-berkeley-launch-rapid-reviews-covid-19-0629

43. Markus A, Oransky I, Retraction Watch. Eye for Manipulation: A Profile of Elisabeth Bik. In: The Scientist Magazine® [Internet]. 7 May 2019 [cited 21 May 2020]. Available: https://www.the-scientist.com/news-opinion/eye-for-manipulation--a-profile-of-elisabeth-bik-65839

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

44. Carneiro CFD, Queiroz VGS, Moulin TC, Carvalho CAM, Haas CB, Rayêe D, et al. Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature. bioRxiv. 2020; 581892. doi:10.1101/581892

45. Fraser N, Brierley L, Dey G, Polka JK, Palfy M, Coates JA. Preprinting a pandemic: the role of preprints in the COVID-19 pandemic. bioRxiv. 2020; 2020.05.22.111294. doi:10.1101/2020.05.22.111294

46. Wickham H, RStudio. rvest: Easily Harvest (Scrape) Web Pages. 2019. Available: https://CRAN.R-project.org/package=rvest

47. Abdill RJ, Adamowicz EM, Blekhman R. International authorship and collaboration in bioRxiv preprints. bioRxiv. 2020; 2020.04.25.060756. doi:10.1101/2020.04.25.060756

48. Chamberlain S, Zhu H, Jahn N, Boettiger C, Ram K. rcrossref: Client for Various “CrossRef” “APIs.” 2020. Available: https://CRAN.R-project.org/package=rcrossref

49. Jahn N, rOpenSci TS (Tuija S reviewed the package for, https://github.com/ropensci/onboarding/issues/115) see, rOpenSci RM (Ross M reviewed the package for, https://github.com/ropensci/onboarding/issues/115) see. roadoi: Find Free Versions of Scholarly Publications via Unpaywall. 2019. Available: https://CRAN.R-project.org/package=roadoi

50. Kearney M. rtweet: Collecting and analyzing Twitter data. J Open Source Softw. 2019;4: 1829. doi:10.21105/joss.01829

51. Steve Royle. Screenager: screening times at bioRxiv. In: quantixed [Internet]. 30 Mar 2020 [cited 22 May 2020]. Available: https://quantixed.org/2020/03/30/screenager-screening-times-at-biorxiv/

52. Venables WN, Ripley BD. Modern Applied Statistics with S. 4th ed. New York: Springer-Verlag; 2002. doi:10.1007/978-0-387-21706-2

53. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biom J Biom Z. 2008;50: 346–363. doi:10.1002/bimj.200810425

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figures

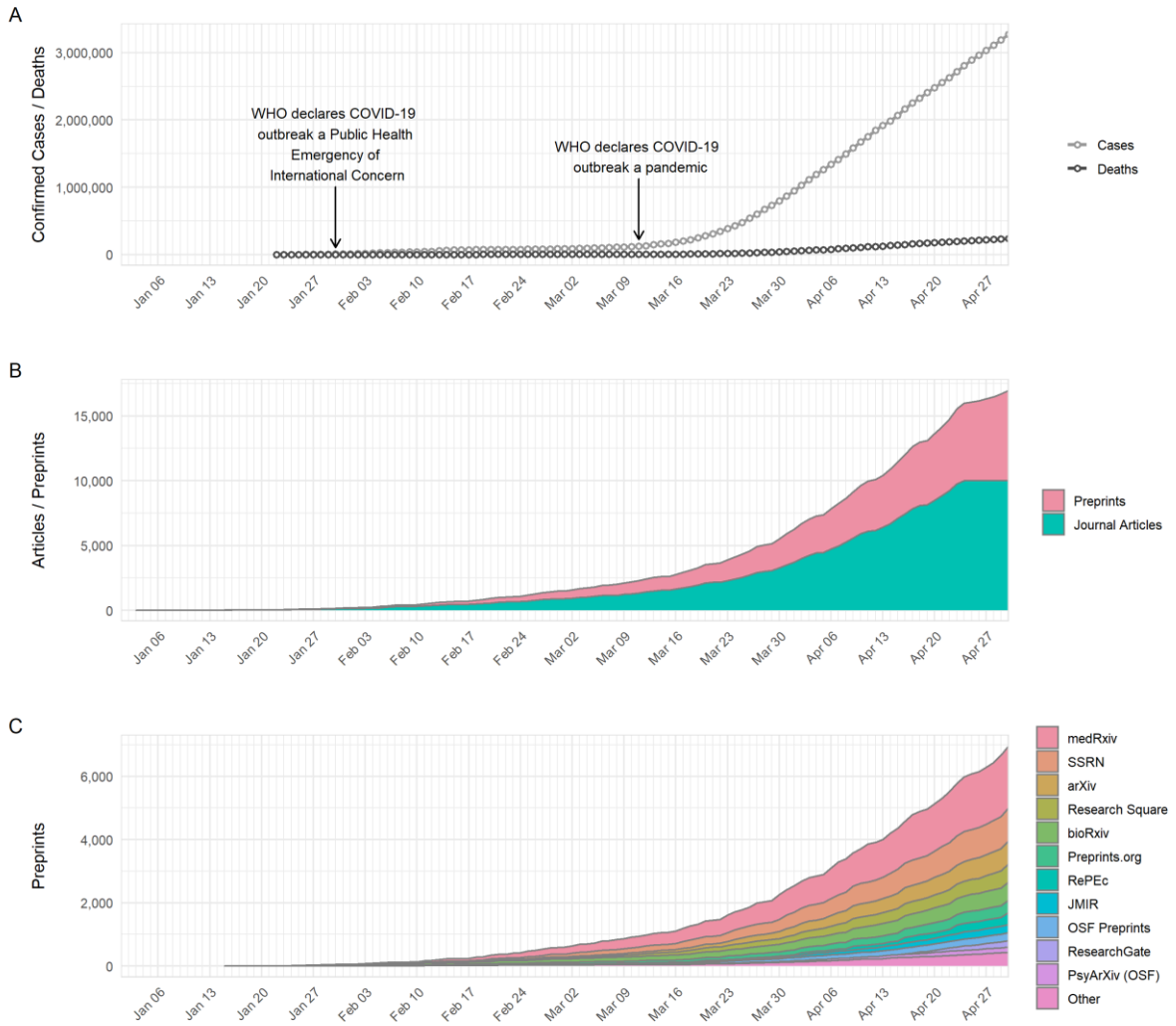


Figure 1. Development of COVID-19 and publication response between January 2020 and April 2020. (A) Number of COVID-19 confirmed cases and reported deaths. Data is sourced from https://github.com/datasets/covid-19/, based on case and death data aggregated by the Johns Hopkins University Center for Systems Science and Engineering (https://systems.jhu.edu/). (B) Cumulative growth of journal articles and preprints containing COVID-19 related search terms. (C) Cumulative growth of preprints containing COVID-19 related search terms, broken down by individual preprint server. Journal data in (B) is based upon data extracted from Dimensions (https://www.dimensions.ai), preprint data in (B) and (C) is based upon data gathered by Fraser and Kramer

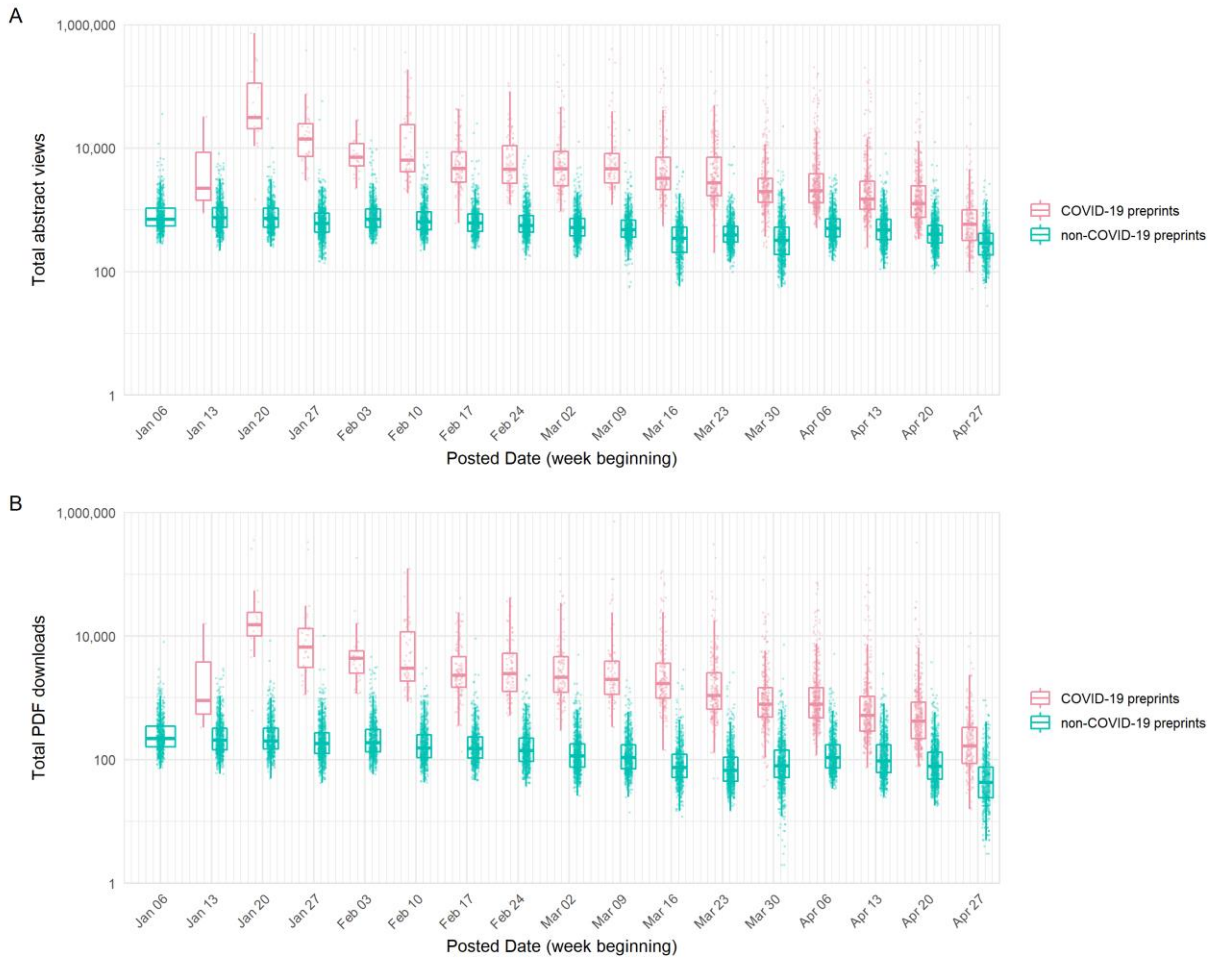
bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 2. Attributes of COVID-19 and non-COVID-19 preprints deposited on bioRxiv and medRxiv between January and April 2020. (A) Number of preprints deposited per week. (B) Preprint screening time. (C) Number of authors per preprint. (D) Proportional representation of 15 most common countries among total corresponding authors of preprints between January and April 2020, stratified by author status where dark fill represents authors previously submitting at least one preprint prior to the COVID-19 pandemic, light fill represents authors submitting a preprint for the first time during the COVID-19 pandemic. (E) License type chosen by authors. (F) Number of versions per preprint. (G) Word counts per preprint. (H) Reference counts per preprint. (I) Percentage of preprints published in peer-reviewed journals (until the end of April 2020). (J) Time taken from posting a preprint until subsequent journal publication for COVID-19 preprints (red), non-COVID-19

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

preprints posted between January - April 2020 (green) and non-COVID-19 preprints posted between September – December 2019 (grey).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



768

Figure 3. Distribution of access statistics for COVID-19 and non-COVID-19 preprints posted on bioRxiv and medRxiv. (A) Total abstract views. (B) Total PDF downloads.

769

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

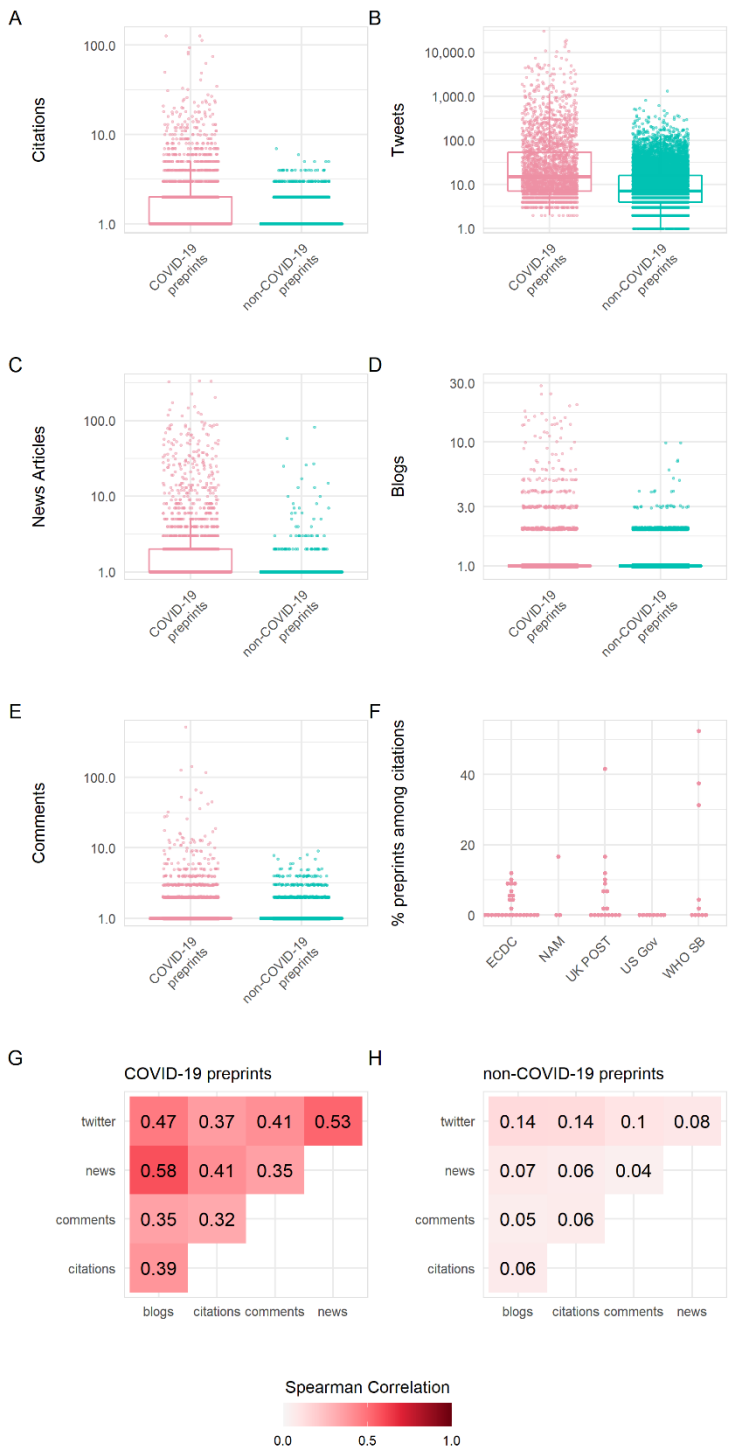


Figure 4. Comparison of citations, tweets, mentions in news articles and blogs for COVID-19 and non-COVID-19 preprints posted on bioRxiv and medRxiv between January and April 2020. (A) Citations per preprint. (B) Tweets per preprint. (C) News article mentions per preprint. (D) Blog mentions per preprint. (E) Number of comments posted on bioRxiv and medRxiv commenting sections per preprint. (F) Percentage of citations made to preprints amongst policy documents from selected sources (ECDC = European Centre for Disease Prevention and Control, NAM = National Academy of Medicine, UK POST = United Kingdom Parliamentary Office of Science and Technology, US HSSCC = United States House Select Subcommittee on the Coronavirus Crisis, WHO SB = World Health Organisation Scientific Briefs). (G) Spearman's correlation matrix between all usage indicators (excluding citations in policy documents) for COVID-19 preprints. (H) Spearman's correlation matrix

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.22.111294; this version posted September 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

between all usage indicators (excluding citations in policy documents) for non-COVID-19 preprints. (A-E) are displayed with log scales, with +1 added for visualisation. Boxplot horizontal lines denote lower quartile, median, upper quartile, with whiskers extending to 1.5*IQR. All boxplots additionally show raw data values for individual preprints with added horizontal jitter for visibility.

786 ## Tables

Table 1. Top 10 cited COVID-19 preprints

Rank	Source	doi	Title	Posted date	Citations
1	biorxiv	10.1101/2020.02.07.937862	Severe acute respiratory syndrome-related coronavirus - The species and its viruses, a statement of the Coronavirus Study Group	11/02/2020	127
2	medrxiv	10.1101/2020.02.06.20020974	Clinical characteristics of 2019 novel coronavirus infection in China	09/02/2020	126
3	medrxiv	10.1101/2020.01.23.20018549	Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions	24/01/2020	112
4	biorxiv	10.1101/2020.01.22.914952	Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin	23/01/2020	93
5	biorxiv	10.1101/2020.01.26.919985	Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov	26/01/2020	83
6	biorxiv	10.1101/2020.01.31.929042	The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells	31/01/2020	79
7	biorxiv	10.1101/2020.01.30.927806	The digestive system is a potential route of 2019-nCov infection: a bioinformatics analysis based on single-cell transcriptomes	31/01/2020	74
8	medrxiv	10.1101/2020.02.10.20021675	Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China	11/02/2020	62
9	biorxiv	10.1101/2020.02.03.931766	Specific ACE2 Expression in Cholangiocytes May Cause Liver Damage After 2019-nCoV Infection	04/02/2020	49
10	medrxiv	10.1101/2020.03.03.20028423	Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts	04/03/2020	48

787

788

Table 2. Top 10 tweeted COVID-19 preprints

Rank	Source	doi	Title	Posted date	Tweets	News articles	Blogs
1	medrxiv	10.1101/2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	17/04/2020	29984	328	24
2	biorxiv	10.1101/2020.01.30.927871	Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag	31/01/2020	18587	92	17
3	medrxiv	10.1101/2020.04.04.20053058	Indoor transmission of SARS-CoV-2	07/04/2020	17494	67	9
4	medrxiv	10.1101/2020.03.22.20040758	Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial	30/03/2020	15337	117	15
5	medrxiv	10.1101/2020.03.09.20033217	Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1	10/03/2020	13407	333	27
6	biorxiv	10.1101/2020.03.13.990226	Reinfection could not occur in SARS-CoV-2 infected rhesus macaques	14/03/2020	10870	225	19
7	medrxiv	10.1101/2020.04.16.20065920	Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19	21/04/2020	10512	329	15
8	medrxiv	10.1101/2020.03.30.20048165	Association of BCG vaccination policy with prevalence and mortality of COVID-19	06/04/2020	10435	3	0
9	medrxiv	10.1101/2020.03.17.20037713	A serological assay to detect SARS-CoV-2 seroconversion in humans	18/03/2020	8094	153	13
10	medrxiv	10.1101/2020.03.24.20042937	Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study	28/03/2020	7427	77	5

789

Table 3. Top 10 COVID-19 preprints covered by news organisations

Rank	Source	doi	Title	Posted date	Tweets	News articles	Blogs
1	medrxiv	10.1101/2020.03.09.20033217	Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1	10/03/2020	13407	333	27
2	medrxiv	10.1101/2020.04.16.20065920	Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19	21/04/2020	10512	329	15
3	medrxiv	10.1101/2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	17/04/2020	29984	328	24
4	biorxiv	10.1101/2020.03.13.990226	Reinfection could not occur in SARS-CoV-2 infected rhesus macaques	14/03/2020	10870	225	19
5	biorxiv	10.1101/2020.03.30.015347	Susceptibility of ferrets, cats, dogs, and different domestic animals to SARS-coronavirus-2	31/03/2020	4399	201	24
6	medrxiv	10.1101/2020.03.23.20039446	Transmission Potential of SARS-CoV-2 in Viral Shedding Observed at the University of Nebraska Medical Center	26/03/2020	4460	172	13
7	medrxiv	10.1101/2020.03.17.20037713	A serological assay to detect SARS-CoV-2 seroconversion in humans	18/03/2020	8094	153	13
8	medrxiv	10.1101/2020.04.07.20056424	Chloroquine diphosphate in two different dosages as adjunctive therapy of hospitalized patients with severe respiratory syndrome in the context of coronavirus (SARS-CoV-2) infection: Preliminary safety results of a randomized, double-blinded, phase IIb clinical trial (CloroCovid-19 Study)	11/04/2020	4503	146	15
9	biorxiv	10.1101/2020.03.08.982637	Aerodynamic Characteristics and RNA Concentration of SARS-CoV-2 Aerosol in Wuhan Hospitals during COVID-19 Outbreak	10/03/2020	972	138	12
10	medrxiv	10.1101/2020.03.11.20031096	Relationship between the ABO Blood Group and the COVID-19 Susceptibility	16/03/2020	3963	127	13

790

Table 4. Top 10 commented on COVID-19 preprints

Rank	source	doi	title	posted date	comments count
1	medrxiv	10.1101/2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	17/04/2020	508
2	medrxiv	10.1101/2020.03.24.20042937	Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study	28/03/2020	141
3	biorxiv	10.1101/2020.01.30.927871	Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag	31/01/2020	127
4	medrxiv	10.1101/2020.04.16.20065920	Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19	21/04/2020	114
5	medrxiv	10.1101/2020.03.11.20031096	Relationship between the ABO Blood Group and the COVID-19 Susceptibility	16/03/2020	66
6	medrxiv	10.1101/2020.03.27.20043752	Forecasting COVID-19 impact on hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months	30/03/2020	61
7	medrxiv	10.1101/2020.03.22.20040758	Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial	30/03/2020	53
8	medrxiv	10.1101/2020.04.05.20054361	Population-level COVID-19 mortality risk for non-elderly individuals overall and for non-elderly individuals without underlying diseases in pandemic epicenters	08/04/2020	47
9	biorxiv	10.1101/2020.01.26.919985	Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov	26/01/2020	44
10	medrxiv	10.1101/2020.03.09.20033217	Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1	10/03/2020	41

791

Table 5. Top 10 most blogged COVID-19 preprints

Rank	Source	doi	Title	Posted date	Tweets	News articles	Blogs
1	medrxiv	10.1101/2020.03.09.20033217	Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1	10/03/2020	13407	333	27
2	biorxiv	10.1101/2020.03.30.015347	Susceptibility of ferrets, cats, dogs, and different domestic animals to SARS-coronavirus-2	31/03/2020	4399	201	24
3	medrxiv	10.1101/2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	17/04/2020	29984	328	24
4	medrxiv	10.1101/2020.03.05.20030502	Clinical presentation and virological assessment of hospitalized cases of coronavirus disease 2019 in a travel-associated transmission cluster	08/03/2020	2675	121	19
5	biorxiv	10.1101/2020.03.13.990226	Reinfection could not occur in SARS-CoV-2 infected rhesus macaques	14/03/2020	10870	225	19
6	biorxiv	10.1101/2020.01.30.927871	Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag	31/01/2020	18587	92	17
7	biorxiv	10.1101/2020.04.01.021196	SARS-CoV-2 neutralizing serum antibodies in cats: a serological investigation	03/04/2020	1298	82	16
8	medrxiv	10.1101/2020.01.23.20018549	Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions	24/01/2020	3278	76	15
9	medrxiv	10.1101/2020.03.22.20040758	Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial	30/03/2020	15337	117	15
10	medrxiv	10.1101/2020.04.07.20056424	Chloroquine diphosphate in two different dosages as adjunctive therapy of hospitalized patients with severe respiratory syndrome in the context of coronavirus (SARS-CoV-2) infection: Preliminary safety results of a randomized, double-blinded, phase IIb clinical trial (CloroCovid-19 Study)	11/04/2020	4503	146	15

792