

RESEARCH ARTICLE

Characteristics of academic publications, preprints, and registered clinical trials on the COVID-19 pandemic

Silvia Gianola¹, Tiago S. Jesus², Silvia Barger¹, Greta Castellini^{1*}

1 Unit of Clinical Epidemiology, IRCCS Istituto Ortopedico Galeazzi, Milan, Italy, **2** Global Health and Tropical Medicine (GHTM), WHO Collaborating Centre for Health Workforce Policy and Planning, Institute of Hygiene and Tropical Medicine—NOVA University of Lisbon, Lisbon, Portugal

* greta.castellini@grupposandonato.it**OPEN ACCESS**

Citation: Gianola S, Jesus TS, Barger S, Castellini G (2020) Characteristics of academic publications, preprints, and registered clinical trials on the COVID-19 pandemic. PLoS ONE 15(10): e0240123. <https://doi.org/10.1371/journal.pone.0240123>

Editor: Tim Mathes, Universitat Witten/Herdecke, GERMANY

Received: June 15, 2020

Accepted: September 21, 2020

Published: October 6, 2020

Copyright: © 2020 Gianola et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data and code used in this study are available on Open Science Framework (osf.io/vby7x).

Funding: SG and GC were supported by the Italian Ministry of Health “Epidemiologia degli studi clinici: qualità delle evidenze e caratteristiche di reporting” (L2060). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

The COVID-19 pandemic has unleashed a deluge of publications. For this cross-sectional study we compared the amount and reporting characteristics of COVID-19-related academic articles and preprints and the number of ongoing clinical trials and systematic reviews. To do this, we searched the PubMed database of citations and abstracts for published life science journals by using appropriate combinations of medical subject headings (MeSH terms), and the COVID-19 section of the MedRxiv and BioRxiv archives up to 20 May 2020 (21 weeks). In addition, we searched Clinicaltrial.gov, Chinese Clinical Trial Registry, EU Clinical Trials Register, and 15 other trial registers, as well as PROSPERO, the international prospective register of systematic reviews. The characteristics of each publication were extracted. Regression analyses and Z tests were used to detect publication trends and their relative proportions. A total of 3635 academic publications and 3805 preprints were retrieved. Only 8.6% ($n = 329$) of the preprints were already published in indexed journals. The number of academic and preprint publications increased significantly over time ($p < 0.001$). Case reports (6% academic vs 0.9% preprints; $p < 0.001$) and letters (17.4% academic vs 0.5% preprints; $p < 0.001$) accounted for a greater share of academic compared to preprint publications. Differently, randomized controlled trials (0.22% vs 0.63%; $p < 0.001$) and systematic reviews (0.08% vs 5%) made up a greater share of the preprints. The relative proportion of clinical studies registered at Clinicaltrials.gov, Chinese Clinical Trial Registry, and EU Clinical Trials Register was 57.9%, 49.5%, and 98.9%, respectively, most of which were still “recruiting”. PROSPERO listed 962 systematic review protocols. Preprints were slightly more prevalent than academic articles but both were increasing in number. The void left by the lack of primary studies was filled by an outpour of immediate opinions (i.e., letters to the editor) published in PubMed-indexed journals. Summarizing, preprints have gained traction as a publishing response to the demand for prompt access to empirical, albeit not peer-reviewed, findings during the present pandemic.

Abbreviations: COVID-19, Coronavirus disease 19; SARS-CoV-2, Severe Acute Respiratory Syndrome coronavirus 2; PROSPERO, international prospective register of systematic reviews; MeSH, Medical Subject Headings; RCTs, randomized controlled trials.

Introduction

The pandemic spread of a novel human coronavirus, termed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses [1] since late 2019 [2], has created a global health emergency. The urgent need to keep the public health response to COVID-19 informed by near-real-time monitoring and assessment [3] has brought with it the expansion of scientific effort and the production of new knowledge at an unprecedented rate [4]. There has been a parallel surge in the research being conducted, posted, and shared in the print and the digital media [5, 6]. The research progress in any area can be quantitatively estimated by bibliometric and scientometric analysis of publications that provides a comprehensive overview of trends in scientific research. Such analyses have been widely used for mapping scientific knowledge in diverse disciplines [7, 8]. For example, to better organize emerging and rapidly scientific developments, scoping reviews coupled with bibliometric and scientometric analysis have documented the networks of contributing authors, institutions, and countries [4, 6, 9, 10]. However, no reviews to date have focused on the epidemiology and the reporting characteristics of publications related to the COVID-19 pandemic [11, 12] or the advances in research during the pandemic [13–15].

Scientific information is distributed through various print and digital channels, among which are the formal publication sources such as registers of primary studies (i.e., trials) and systematic reviews, preprints, and academic publications. A clinical trials registry is a platform in which clinical trials are registered and cataloged. Clinical trials play an important role in response to the COVID-19 pandemic as a means to translate evidence from basic research into clinical practice [16, 17]. A systematic review protocol registry is an international database in which prospective systematic reviews in health and social care are listed. Systematic reviews should be registered at inception (at the protocol stage) to help avoid outcome reporting biases, publication bias, unplanned duplication and waste of resources, especially during the present global health emergency [18, 19]. Preprints are preliminary reports of work and have not yet undergone peer review. As such, they should not be relied on to guide clinical practice or health-related behavior and should not be reported in the news media as established information [20]. Preprint servers are open access online repositories that house preprint research articles by authors who choose to make their research immediately and freely available to the public. Research articles can receive commentary and peer review prior to journal submission, thus accelerating the dissemination of scientific findings [21, 22]. Preprints posted during the Ebola and Zika outbreaks divulged novel analyses and new data, while many of those that later appeared in academic publications were available for more than 100 days before publication [21]. Many preprints never undergo peer review; for instance, less than 5% of journal articles about the Ebola and the Zika virus were posted as preprints prior to publication in journals [21]. The information in preprints lacks the scrutiny and validity of an external, scientific review [23]. Nevertheless, analysis of preprints and ongoing investigations (trials and reviews) is essential for scientific advancement where timeliness is key, though peer reviews are still expedited for new relevant research.

With this context in mind, we wanted to take a snapshot of the amount of scientific development devoted to COVID-19 and gain an idea of the current status of global research. We also thought it necessary to obtain an insight into the reporting of future research, its amount, design, publication venues, and characteristics. To do this, we:

1. Compared the amount of COVID-19 related research in academic publications and preprints (not peer reviewed) stratified by key characteristics: species (*Humans, Animals*), study design (e.g., systematic review, randomized controlled trials—RCTs), and research area (e.g., drug treatment)

2. Evaluated the amount and key characteristics (e.g., research area) of ongoing COVID-19 research contained in clinical trials registers and systematic reviews.

Methods

For this cross-sectional study, COVID-19 related-publications listed in electronic databases as of 20 May 2020 were reviewed.

Data sources

We searched the PubMed database for COVID-19-related academic publications through its indexation system. PubMed is a comprehensive database that comprises more than 30 million citations (as of May 2020) for biomedical literature from MEDLINE, life science journals, and online books [24]. There is evidence in the health field that adding databases other than PubMed has only modest impact on the results of systematic reviews [25]. The medical subject headings (MeSH terms) in PubMed are organized in a hierarchical tree and assigned to each paper by subject-specialist indexers. For searching COVID-19-related articles, we used the supplementary concept records (SCRs). MEDLINE indexers regularly come across substances in the literature that are not currently MeSH headings. When this happens, the National Library of Medicine staff adds these substances to the MeSH vocabulary as SCRs. While MeSH headings are updated annually, new SCRs are added weekly. Therefore, COVID-19 articles are systematically indexed by research topic regardless of the specific words used by the authors. We conducted target searches in PubMed that included the use of appropriate combinations of MeSH terms and SCRs related to COVID-19 and related terms (e.g., SARS-CoV-2, coronavirus disease), filtered by species, research design, and research area ([S1 Appendix. PubMed database](#)).

We searched two popular sources for COVID-19-related preprints [26], the MedRxiv and BioRxiv databases, here defined as proxy indicators of published empirical not yet peer-reviewed literature. Articles on these preprints servers are examined by in-house staff who check for such issues as plagiarism and incompleteness [27] ([S1 Appendix. Preprints MedRxiv and BioRxiv databases](#)).

We then investigated all primary registries that meet the International Committee of Medical Journal Editors (ICMJE) criteria according to the WHO Registry Network. Primary registries in the WHO Registry Network meet [specific criteria](#) for content, quality, and validity, accessibility, unique identification, technical capacity and administration [28]. Additionally, we searched the PROSPERO international prospective register of systematic reviews database ([S1 Appendix. Primary register and PROSPERO register](#)).

All searches were run on 20 May 2020 (covering a 21-week period), without language restrictions.

Data extraction

Data records retrieved from PubMed, MedRxiv, BioRxiv, trials registers, and PROSPERO databases were imported to Excel for further grouping and analysis. Data extraction was performed by one author and then checked by the second author. Any uncertainties were discussed between the data extractors.

For academic publications and preprints, we extracted the bibliometric data and then analyzed and summarized the subsets of the study or reporting characteristics: species (*Humans*, *Animals*); publication type (any type of article, systematic review, RCTs, epidemiologic studies, Letter to the Editor with or without original data); the distribution by research area (e.g., vaccine, drug treatment, rehabilitation, diagnostic testing, measures for infection prevention and

control including social distance, masks); and biology according to the indexation facilities and the respective search tags or filters. Since the preprint servers have no such filters/tags, we manually extracted and coded the topics. Study designs such as laboratory experiments (e.g., genomics) and prognostic models were included in the epidemiology studies.

For the registers listing clinical trials, the following characteristics were filtered/tagged and extracted: country, study type, phase, recruiting status. Species and research area/tag were extracted from the records retrieved from the PROSPERO database.

Data analysis

Descriptive statistics were initially used for analysis. The total number of publications for each data element (e.g., publication type), percentages, and frequencies were computed and displayed in tabular format, paired by academic and preprint sources. The total number of publications over time (weekly) was plotted onto a run chart. Simple linear regression with analysis of variance (ANOVA) was applied to analyze the growth in the number of COVID-19-related publications over time. Finally, two tailed two-sample Z-tests were performed to determine whether the proportion of each characteristic of the academic journal articles (indexed in PubMed) differed significantly from those contained in the preprint sources. Statistical significance was set at <0.05 . Data analysis was performed using Stata Statistical Software. Release 16. (StataCorp LP, College Station, TX, USA).

Results

1. Amount and characteristics of COVID-19-related research reported in academic and preprint (not peer reviewed) publications

From December 2019 through May 2020, a total of 7440 COVID-19-related publications were retrieved: 3635 (48.9%) academic publications and 3805 (51.1%) preprints. Fig 1 presents the number of academic and preprint publications by week from December 2019 through May

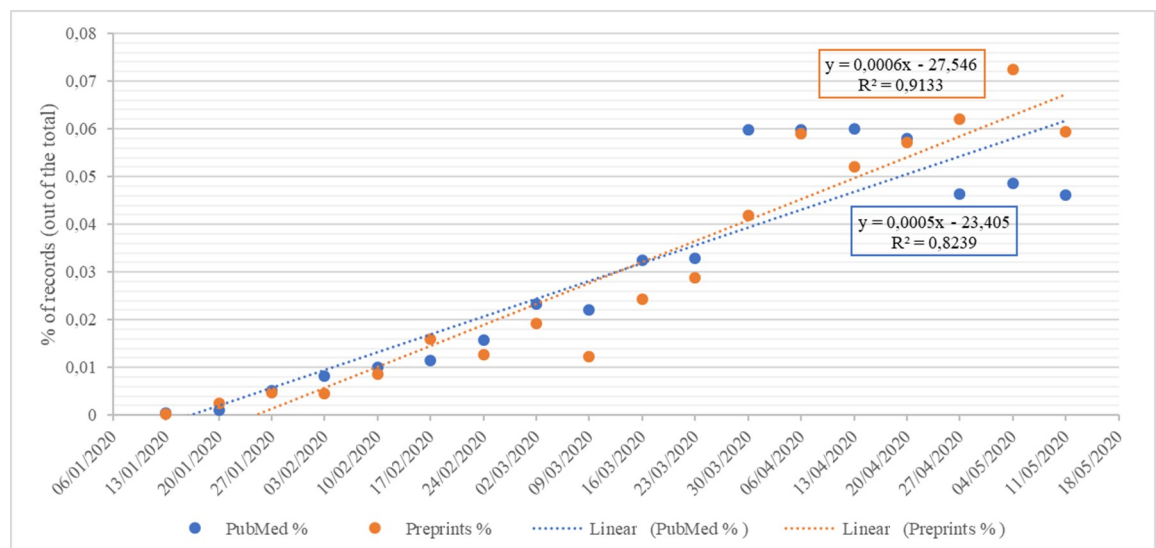


Fig 1. Linear regression over time for academic and preprint publications. Dates: 10 January 2020—first genomic data of its etiological agent (SARS-CoV-2); 20 January 2020—reports of confirmed cases from three countries outside China; 24 January 2020—first European case; 30 January 2020—the WHO declares the novel coronavirus outbreak (2019-nCoV) a Public Health Emergency of International Concern (PHEIC); 11 March 2020—the WHO declares a status of pandemic; 25 March 2020—cases confirmed in all European Union/European Economic Area (EU/EEA) countries and in more than 150 countries worldwide.

<https://doi.org/10.1371/journal.pone.0240123.g001>

2020. The increase in the number of academic and preprint publications was statistically significant for both types (academic $p < 0.001$, $r^2 = 0.8239$; preprints $p < 0.001$, $r^2 = 0.9133$). **Figs 2 and 3** present the absolute and the relative frequency of publications by week. **Fig 4** presents the trend of the ratio between the relative frequency of academic and preprint publications by week. There was no statistically significant difference in the rate of weekly increase in relative frequency between the two groups ($p = 0.2388$, $r^2 = 0.0856$).

Table 1 presents the reporting characteristics of academic and preprint COVID-19-related publications. In both types of publications, human subjects were predominant; the most mentioned research areas were prevention and control (26.1%, 950/3635 academic vs 42.4%, 1615/3805 preprints; $p < 0.001$) and diagnosis (21.5%, 781/3635 academic vs 25.3%, 962/3805 preprints; $p < 0.001$).

There were statistically significant differences in the relative frequency of study types between academic and preprint publications. For instance, RCTs were significantly more frequently reported in preprints than in academic publications (academic 0.2%, 8/3635 vs. preprints 0.6%, 24/3805; $p < 0.001$). Similarly, systematic reviews appeared significantly more often in preprints (academic 0.8%, 29/3635 vs. preprints 5.0%, 193/3805; $p < 0.001$). Observational studies were significantly more common in preprint than in academic publications (academic 0.6%, 21/3635 vs. preprints 24.7%, 940/3805; $p < 0.001$). In contrast, case reports and letters to the editor were significantly more common in academic publications (case reports in academic publications 6.0%, 219/3635 vs. preprints 0.9%, 35/3805; $p < 0.001$; letters to the editor in academic publications 17.4%, 632/3635 vs. preprints 0.5%, 19/3805; $p < 0.001$). Overall, in both types of publications the most often mentioned research areas were prevention and control (26.1%, 950/3635 vs. 42.4%, 1615/3805; $p < 0.001$) and diagnosis (21.5%, 781/3635 vs. 25.3%, 962/3805).

Only 8.6% (329/3805) of preprints appeared in academic journals (as of May 2020), with a median of 28 days (interquartile range [IQR] 14–52; min–max: 0–117) from preprint posting

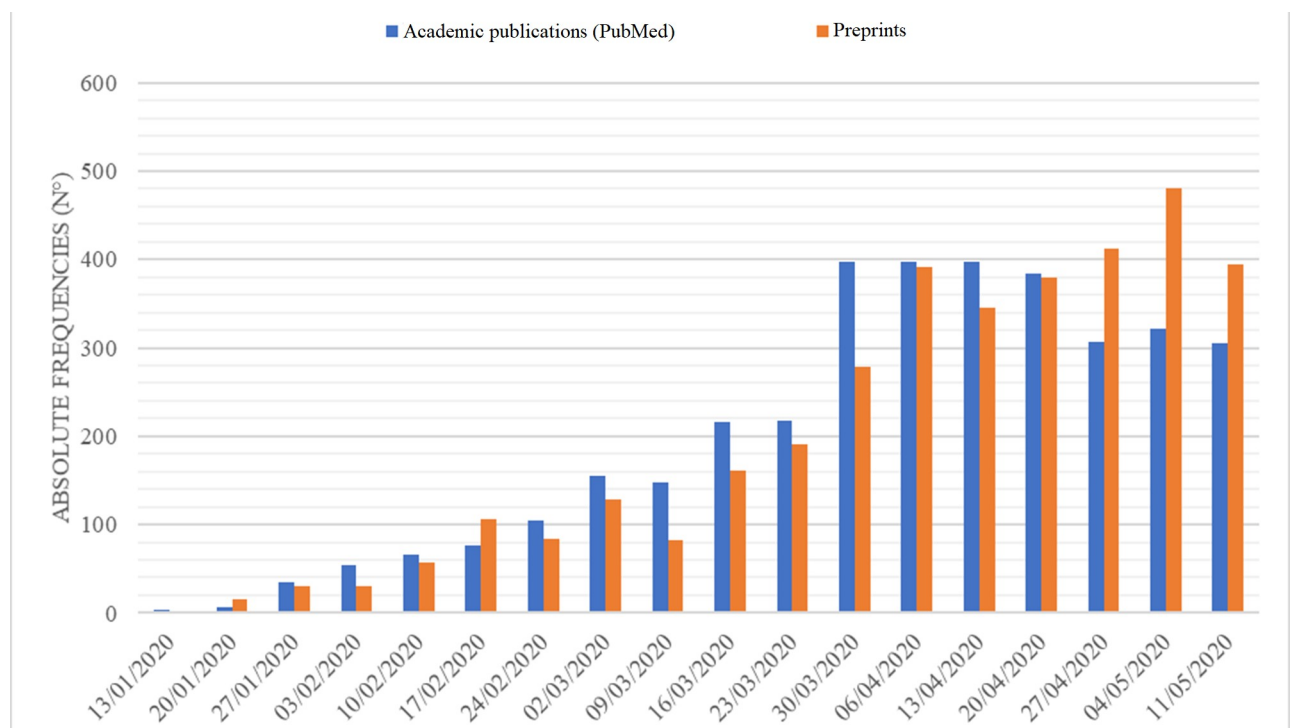


Fig 2. Absolute frequency of academic and preprint publications by week.

<https://doi.org/10.1371/journal.pone.0240123.g002>

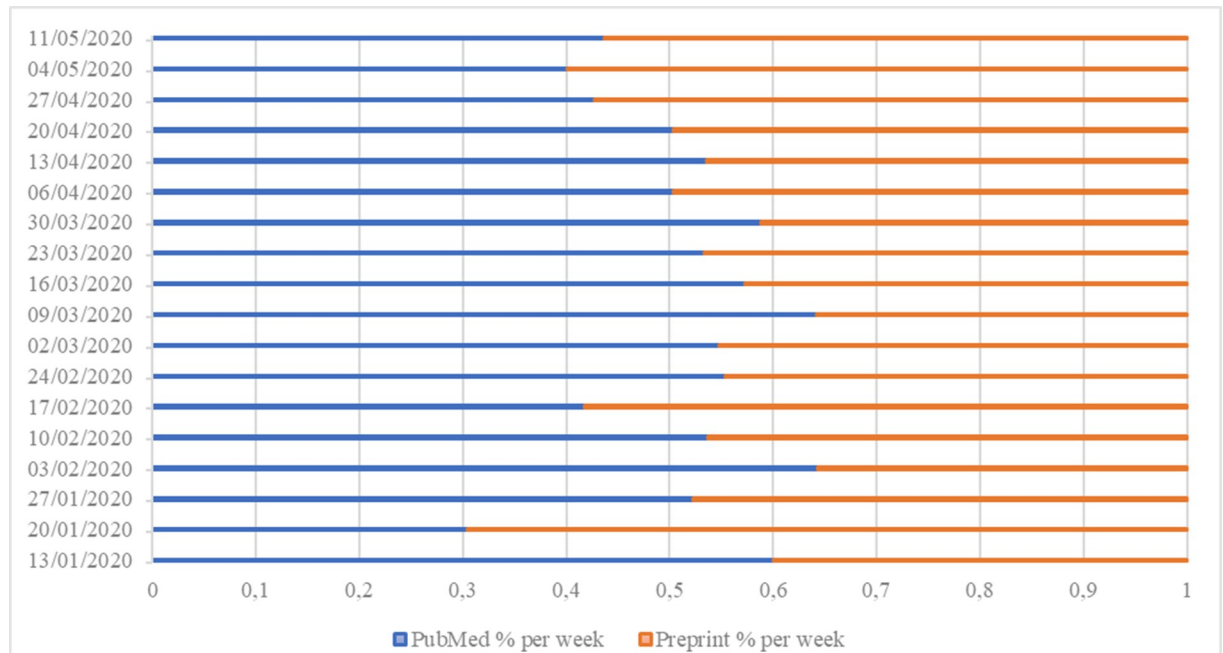


Fig 3. Relative distribution of academic and preprint publications by week.

<https://doi.org/10.1371/journal.pone.0240123.g003>

to publication online. Observational studies (87.8%, n = 289/329) and RCTs (5.6%, n = 18/329) were most often published, while prevention and control (40.7%, n = 134/329) and diagnosis (31.6%, n = 104/329) were the most frequently mentioned research areas.

2. Characteristics of ongoing COVID-19-related research in the registers (trials and systematic reviews)

Table 2 shows that 1621/339,863 (0.5%) records in Clinicaltrials.gov were COVID-19-related and that 652/32553 (2%) records in the Chinese Clinical Trial Registry (ChiCTR) were COVID-19-related. No other register had over 200 records of COVID-19-related research.

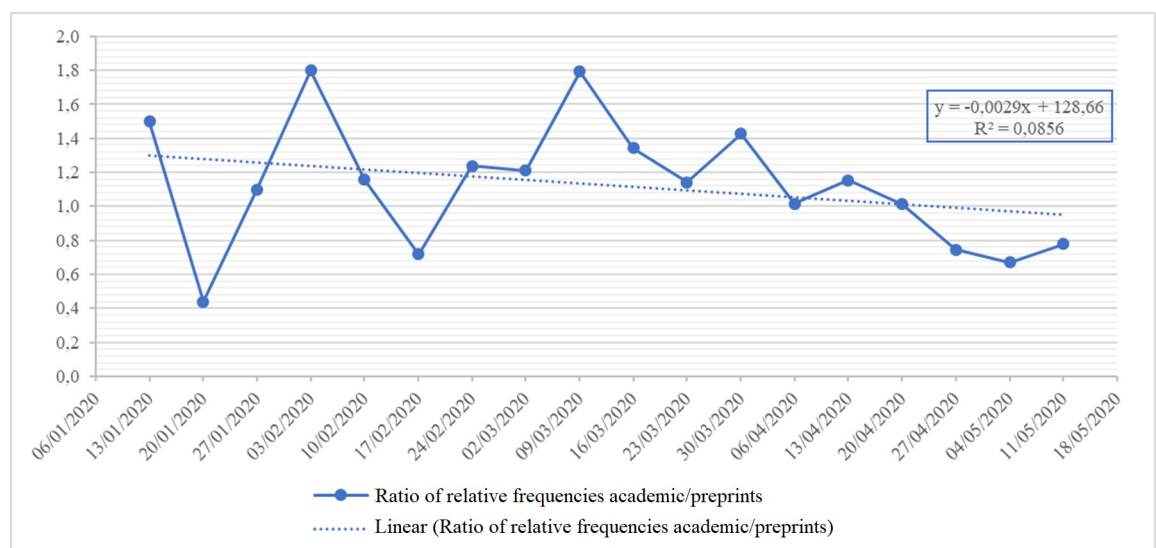


Fig 4. Linear regression of ratios between relative frequency (academic/preprint publications) by week.

<https://doi.org/10.1371/journal.pone.0240123.g004>

Table 1. Characteristics of COVID-19-related academic (PubMed) and preprint publications.

Characteristic		PubMed N = 3635 No. (%)	Preprints N = 3805 No. (%)	Z value	% Difference	P value
Species	<i>Humans</i>	3507 (96.5)	3634 (95.5)	2.14	1.0	<0.05
	<i>Animals</i>	128 (3.5)	93 (2.4)	2.74	1.1	<0.05
Publication type	Systematic review and/or meta-analysis	29 (0.8)	193 (5.0)	-10.83	-4.2	< 0.0001
	Randomized controlled trial	8 (0.2)	24 (0.6)	-2.71	-0.4	<0.05
	<i>Phase I</i>	1	0	1.02	1.0	>0.05
	<i>Phase II</i>	3	3	0.06	0.0	>0.05
	<i>Phase III</i>	0	2	-1.38	-2.0	>0.05
	<i>Phase IV</i>	0	3	-1.69	-3.0	>0.05
	<i>Protocols</i>	0	4	-1.96	-4.0	>0.05
	Epidemiologic studies	222 (6.1)	3522 (92.6)	-74.55	-86.5	< 0.0001
	<i>Observational studies</i>	21 (0.6)	940 (24.7)	-31.02	-24.1	< 0.0001
	Case report	219 (6.0)	35 (0.9)	12.12	5.1	< 0.0001
	Letter to the Editor	632 (17.4)	19 (0.5)	25.77	16.9	< 0.0001
	Research Area/ Tag	Vaccine	101 (2.8)	114 (2.9)	-0.56	-0.1
Drug therapy		57 (1.6)	325 (8.5)	-13.62	-6.9	< 0.0001
Diagnosis		781 (21.5)	962 (25.3)	-3.87	-3.8	< 0.0001
Prevention and control (e.g., masks, social distancing)		950 (26.1)	1615 (42.4)	-14.80	-16.3	< 0.0001
Rehabilitation (e.g., pulmonary rehabilitation)		23 (0.6)	5 (0.1)	3.53	0.5	< 0.0001
Prognosis		123 (3.4)	844 (22.2)	-24.10	-18.8	< 0.0001
Biology/Genetic		43 (1.1)	866 (22.8)	-28.41	-21.7	< 0.0001

Note: the sum of the characteristics does not correspond to the total number of publications since specifications or other may overlap.

<https://doi.org/10.1371/journal.pone.0240123.t001>

Table 3 presents the characteristics of the two primary registries with the highest number of COVID-19-related research studies. The most frequent study types were interventional studies in Clinicaltrials.gov and in ChiCTR (60% and 50%, respectively), with phase 3 the most frequent in Clinicaltrials.gov (15%). The most frequent recruiting status was “ongoing” in all registers (min-max: 47%-99%).

Table 4 presents the characteristics of the records retrieved from the PROSPERO database for COVID-19-related systematic reviews: 962 systematic reviews were registered; almost all focused on humans (99.7%); the two most frequent research areas were treatment (19.1%) and health impact (16.6%).

Discussion

During the first half of 2020 (about 5 months/21 weeks), the spread of COVID-19 matched the rapid and statistically significant increase in the number of publications appearing in academic journals and on preprint servers (3805 and 3635 records, respectively; $p < 0.001$). Clinicians,

Table 2. Primary registries.

	Publications	
	COVID-19-related records No. (%)	Total no. of records
Clinicaltrials.gov (USA)	1561(0.5)	339,863
Chinese Clinical Trial Registry (ChiCTR)	648 (2.0)	32,553
EU Clinical Trials Register (EU-CTR)	196 (0.5)	37,185
Iranian Registry of Clinical Trials (IRCT)	195 (0.8)	24,573
Clinical Trials Registry—India (CTRI)	99 (7.0)	1475
German Clinical Trials Register (DRKS)	67 (1.7)	3959
Japan Primary Registries Network (JPRN)	64 (0.2)	28,794
Australian New Zealand Clinical Trials Registry (ANZCTR)	52 (0.2)	27,187
The Netherlands National Trial Register (NTR)	52 (0.6)	8613
International Standard Randomized Control Number (ISRCTN)	35 (0.4)	11,381
Brazilian Clinical Trials Registry (ReBec)	10 (0.2)	4085
Pan African Clinical Trial Registry (PACTR)	6 (0.5)	1162
Clinical Research Information Service (CRiS), Republic of Korea	2 (0.4)	505
Lebanese Clinical Trials Registry (LBCTR)	2 (3.0)	71
Peruvian Clinical Trial Registry (REPEC)	0 (0.0)	1849
Sri Lanka Clinical Trials Registry (SLCTR)	1 (0.3)	349
Thai Clinical Trials Registry (TCTR)	9 (3.0)	336
Cuban Public Registry of Clinical Trials(RPCEC)	19 (NA)	NA

<https://doi.org/10.1371/journal.pone.0240123.t002>

researchers, public health authorities, and other stakeholders (e.g., government administrators) need to be able to identify in a timely manner accurate and reliable health information from trustworthy data sources. With this study we analyzed the frequency and type of publication covering the initial scientific response to the COVID-19 pandemic as of 20 May 2020. In general, we noted that to fill the void left by the absence of primary studies, academic journals ran opinion pieces (i.e., letters, editorials, commentary), whereas preprint servers listed empirical studies, including epidemiological findings and research results in biology, infection diagnosis and prevention.

Our findings reflect the progressive advances in attempts to answer questions from the scientific community. While the total number of publications rose dramatically in the first months of the study period (not counting the last few weeks because of indexation delays), academic sources provided scarce higher-hierarchy evidence in the form of systematic reviews/meta-analysis ($n = 29$) or randomized controlled trials ($n = 8$). This observation is shared by previous studies that found few high-quality studies on COVID-19 [17, 29, 30]. Since there has been insufficient time to design, approve, conduct, and conclude such studies, reports of preliminary results can be expected over the upcoming months. A previous cross-sectional study on COVID-19-related publications revealed that only a quarter of the first 2118 articles contained original data [31]. In our study, the number of registered trial protocols listed as “ongoing” in Clinicaltrials.gov and in ChiCTR ($n = 795$ and $n = 308$, respectively) signals an upward trend. We retrieved a total of 2273 trials from the two registers and noted a phenomenal growth rate in both. As of 22 February 2020, merely 171 COVID-19-related interventional trials were registered (138 in ChiCTR and 33 in ClinicalTrials.gov) [32]; one month later (24 March 2020) the number was 614 COVID-19-related interventional trials (471 in ChiCTR and

Table 3. Characteristics of primary registries with highest number of COVID-19-related trials. (absolute frequencies).

Characteristic		Clinicaltrials.gov N = 1621 no. (%)	Chinese Clinical Trial Registry (ChiCTR) N = 652 no. (%)	EU Clinical Trials Register N = 196 no. (%)
Country	Africa	56 (3.5)	-	-
	Central America	3 (0.2)	-	-
	East Asia (China)	120 (7.4)	652 (100.0)	-
	<i>Japan</i>	4 (0.2)	-	-
	Europe	628 (38.7)	-	196 (100.0)
	Middle east	83 (5.1)	-	-
	North America	380 (23.4)	-	-
	<i>Canada</i>	44 (2.7)	-	-
	<i>United States</i>	326 (20.1)	-	-
	<i>Mexico</i>	18 (1.1)	-	-
	North Asia	13 (0.8)	-	-
	Pacifica	10 (0.6)	-	-
	South America	48 (2.9)	-	-
	South Asia	20 (1.2)	-	-
	South-East Asia	18 (1.1)	-	-
	Study Type	Interventional	940 (57.9)	323 (49.5)
Observational		663 (40.9)	260 (39.9)	-
Other (i.e., patient registries)		125 (7.7)	69 (10.6)	-
Phase	Early Phase 1	19 (1.2)	218* (33.4)	-
	Phase 1	91 (5.6)	13 (2.0)	6 (3.1)
	Phase 2	375 (23.1)	8** (1.2)	99 (50.5)
	Phase 3	241 (14.9)	3*** (0.5)	75 (38.3)
	Phase 4	55 (3.4)	67 (10.3)	33 (16.8)
	Not applicable	282 (17.4)	212 (32.5)	-
Recruiting status	Recruiting	795 (49.0)	308 (47.2)	194 (98.9)
	Complete	76 (4.7)	53 (8.1)	0
	Suspended/temporarily halted	7 (0.4)	16 (2.5)	0
	Other (e.g. withdrawn)	8 (0.5)	275 (42.2)	1 (0.5)

Na*: records not available because not possible to retrieve.

Data were collected as reported in primary registers.

* phase 0 for ChiCTR.

**3 trials were phase I-II.

***1 trial was phase II-III.

<https://doi.org/10.1371/journal.pone.0240123.t003>

143 in ClinicalTrials.gov) [33]. The rise may be linked to the breakthrough news from China and other countries that announced a successful drug treatment [34]. In retrospect, however, the announcement by Chinese scientists that made claims for chloroquine was not a true breakthrough in COVID-19 treatment; nonetheless, it could have attracted low-quality research and clinical practice [17, 35, 36].

Many of the systematic review protocols registered in PROSPERO (n = 962) are focused on treatment; however, they may encounter limitations due to the scarcity of the number or of the quality of clinical trials [17]. The secondary literature can be supported by indirect evidence from primary studies involving patients different from those of interest (i.e., COVID-19), as suggested by the Grading of Recommendations Assessment, Development and Evaluation

Table 4. Characteristics of COVID-19-related systematic review protocols.

Characteristic		Prospero registration N = 962	
		No.	%
Species	<i>Humans</i>	959	99.0
	<i>Animals</i>	3	0.3
Research Area/Tag as reported in PROSPERO	Chinese medicine	70	7.0
	Diagnosis	52	5.0
	Epidemiological	155	16.0
	Genetics	7	0.7
	Health impacts	160	17.0
	Mental health	76	8.0
	Other	31	3.0
	Personnel protective equipment	17	2.0
	Prognosis	50	5.0
	Public health	10	1.0
	Transmission	26	3.0
	Treatments	184	19.0
	Vaccines	3	0.3

<https://doi.org/10.1371/journal.pone.0240123.t004>

(GRADE) approach [37]. The search for answers, treatments, and vaccines during a pandemic ought to be based on the highest levels of evidence such as systematic reviews and randomized controlled trials. Furthermore, companies and universities have begun to accelerate the discovery of experimental drugs and vaccines for COVID-19 through their research and development pipeline [38]. The World Health Organization (WHO)'s Blueprint list recognizes coronaviruses as infectious agents of priority importance: given the public health emergencies of international concern and the current absence of efficacious drugs and vaccines, research and development need to be accelerated if diseases caused by coronaviruses are to be controlled [39].

For the duration of the COVID-19 pandemic, several scientific journals allow access to publications through preprints posted online [40–42]. However, the credibility of preprints not published in a reputable peer-reviewed journal is uncertain. We found that only 8% of preprints were published in academic journals after undergoing peer review. We wonder whether time constraints on converting preprints into a final peer-reviewed version and then waiting for publication are the reasons for such a low percentage. The time to publication of a definitive manuscript can take up to 166 days for a preprint to be published with its DOI in a journal article [43]. One fourth of the published preprints we retrieved took between 52 and 117 days to appear in an academic journal. Publication setbacks, often due to bottlenecks in the peer-review process, can delay the timely dissemination of information in a global health emergency like the present COVID-19 pandemic [43]. Furthermore, unpublished studies generate a waste of effort and money invested in health and medical research globally [44, 45]. We noted that findings in areas like mathematics were published only on preprint servers like ArXiv [46, 47]. The potential harm from posting erroneous provisional research is one reason why the medical community was so cautious about preprints in the first place [48]. But it is interesting to observe that preprints are now gaining more attention and citations than the corresponding peer-reviewed article (i.e., academic publication) [49]. While it is still soon to analyze citation trends for COVID-19-related articles, the publication trends we observed (as of 20 May 2020) were remarkable for the greater number of preprints than academic publications and the

higher percentage of empirical research reported in preprints. This observation contrasts with the findings by Lv et al. (as of 6 February 2020) who reported that academic journal articles accounted for 77.1% of study publications, preprints for 14.1%, while 8% were published only online [50]. A plausible explanation for these discrepancies is the sheer amount of recent empirical studies more rapidly accessible in preprint databases than in academic journals.

Preprints provide a faster and iterative alternative or complement to journal publication and the peer-review system. In addition, they catalyze biomedical discovery, support career advancement, and improve scientific communication [43]. If all scientific publishers endorsed preprint posting for research, it would serve as a preliminary step to publication and a clear signal to all scientists that preprints are an integral part of scientific communication [21]. In essence, better utilization of preprints would mean better fruition of trusted science. By tracking the preprint along the publication process the reliability of research can be controlled and differentiated from the final peer-reviewed publication. Also, it might also benefit the author and the scientific community alike: the former can share their research and attract immediate attention, while the latter can directly access the research paper, give feedback to improve the manuscript, and open a discussion leading to new ideas, follow-up studies or collaboration with other research groups [51]. This may be viewed as a necessary step toward a more open and transparent peer-review process, which forms the cornerstone of scientific activities and part of a system that ensures that only quality research papers are released into the scientific community [23]. Commentaries, letters to the editor, and other forms of non-research publications were largely found in academic journals. Though less suitable as preprints, they can still benefit from external peer-review since they are biased or present personal perspectives. Informed readers of empirical research can more objectively appraise the research methods, the data presented, the validity of its conclusions or limitations in either preprint or definitive versions.

The challenge in this emergency is to increase the amount of completed and published research and translate it into practice. However, a recent commentary in *International Immunopharmacology* stated, “quantity does not equal quality” [52]. While the dissemination of data is an essential part of the response, the detection of poor methodology and the retraction of clinical trials published in reputable journals are unfortunate examples of misconduct and rushed research in this pandemic [17, 53].

Limitations of the study

This study covers only journals and publications indexed in PubMed, a comprehensive but not exhaustive database for biomedical literature. We searched only a few of the repositories and servers for posting preprints (i.e., bioRxiv and medRxiv). We used the COVID-19 section of these servers to investigate the preprint versions at the time of our data extraction. We collected data within a tightly defined window (December 2019 through May 2020). Since no filter nor topics exist for selecting preprint categories and the average time the servers take to display search results can be long [54], we performed hand screenings. This partly reduces the validity of the comparison of study types (academic vs. preprint). An illustrative example is that of the epidemiologic studies: we retrieved different study types in preprints (e.g., case-control studies, cross-sectional studies, biological and mathematical models), whereas when searching PubMed we used MeSH terms such as clustered study design (e.g., case-control studies, cross-sectional studies). In addition, epidemiological studies are not indexed in PubMed as a “publication type” like RCTs, for instance. This can affect the reliability, accuracy, and speed of indexation of epidemiological studies in PubMed. compared to preprint servers. It may also explain why many studies retrieved from PubMed were not yet classified by study

type (the total sum of the articles indexed by study type was clearly below the total number of papers retrieved). Therefore, clear-cut publication types like RCTs or letters to the editors may have had faster, more reliable indexation. It is reasonable to infer that many of the studies without indexation for a publication or study type could be some format other than an RCT or commentaries/letters (e.g., epidemiologic, secondary data analysis, etc.). Although indexation of COVID-19-related research has been prioritized, delays may apply and these may have been visible in the data we obtained during the weeks of May 2020.

When needed (when the search filters or tags could not be reliably applied), one researcher performed manual coding and the other verified it rather than involving two independent researchers. This option is typical of rapid review types and a compromise to expedite the results for which timeliness is key, perhaps at the expense of a certain degree of uncertainty about the accuracy of the results. We did not adjust the characteristics for a country's population size, for example using the World Bank's data to facilitate comparison of the number of studies per country (<https://data.worldbank.org/>).

Conclusion

The number of COVID-19-related publications has increased dramatically since the outbreak began in late 2019. The bulk of the scientific literature in the first five months of the epidemic comprised short reports of data, opinions or perspectives (e.g., letters to the editor or commentaries) or reported (secondary) data rather than primary data. Numerous clinical trials and systematic reviews are registered but their results have not yet been published. The focus of research is effective preventive/control and therapeutic strategies, reflecting the initial response by the research community. Preprints can provide readers with direct, rapid access to research findings. Despite the deluge of papers [55], screening and reviewing preprints might benefit both the authors and the scientific community, while informing the general public with timely access to scientific content and discussions that could boost scientific credibility [41].

In brief, we noted a huge increase in the amount of scientific papers within a short period of time: COVID-19 has aroused an unprecedented interest in research, a demand for answers, and rapid publication of findings. Nevertheless, efforts and time should be devoted to scientifically sound information rather than rushing to publish results that may not be fit to print. Overproduction of research is another form of waste that undermines public trust in science [56].

Supporting information

S1 Appendix.
(DOCX)

Author Contributions

Conceptualization: Silvia Gianola, Tiago S. Jesus, Greta Castellini.

Data curation: Silvia Gianola, Greta Castellini.

Formal analysis: Silvia Gianola, Greta Castellini.

Investigation: Silvia Gianola, Silvia Barger, Greta Castellini.

Methodology: Silvia Gianola, Tiago S. Jesus, Greta Castellini.

Project administration: Silvia Gianola, Greta Castellini.

Writing – original draft: Silvia Gianola, Greta Castellini.

Writing – review & editing: Silvia Gianola, Tiago S. Jesus, Silvia Barger, Greta Castellini.

References

1. 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(4):536–44. <https://doi.org/10.1038/s41564-020-0695-z> PMID: 32123347
2. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International journal of antimicrobial agents.* 2020; 55(3):105924.
3. Cowling BJ, Leung GM. Epidemiological research priorities for public health control of the ongoing global novel coronavirus (2019-nCoV) outbreak. *Euro surveillance: bulletin European sur les maladies transmissibles = European communicable disease bulletin.* 2020; 25(6).
4. Haghani M, Bliemer MCJ, Goerlandt F, Li J. The scientific literature on Coronaviruses, COVID-19 and its associated safety-related research dimensions: A scientometric analysis and scoping review. *Safety science.* 2020:104806.
5. Odone A, Salvati S, Bellini L, Bucci D, Capraro M, Gaetti G, et al. The runaway science: a bibliometric analysis of the COVID-19 scientific literature. *Acta Biomed.* 2020; 91(9-S):34–9. <https://doi.org/10.23750/abm.v91i9-S.10121> PMID: 32701915
6. Nowakowska J, Sobocinska J, Lewicki M, Lemanska Z, Rzymiski P. When science goes viral: The research response during three months of the COVID-19 outbreak. *Biomed Pharmacother.* 2020; 129:110451. <https://doi.org/10.1016/j.biopha.2020.110451> PMID: 32603887
7. Falagas ME, Papastamataki PA, Bliziotis IA. A bibliometric analysis of research productivity in Parasitology by different world regions during a 9-year period (1995–2003). *BMC infectious diseases.* 2006; 6:56. <https://doi.org/10.1186/1471-2334-6-56> PMID: 16545114
8. Zyoud SH. Global research trends of Middle East respiratory syndrome coronavirus: a bibliometric analysis. *BMC infectious diseases.* 2016; 16:255. <https://doi.org/10.1186/s12879-016-1600-5> PMID: 27267256
9. Bonilla-Aldana DK, Quintero-Rada K, Montoya-Posada JP, Ramirez-Ocampo S, Paniz-Mondolfi A, Rabaan AA, et al. SARS-CoV, MERS-CoV and now the 2019-novel CoV: Have we investigated enough about coronaviruses?—A bibliometric analysis. *Travel medicine and infectious disease.* 2020; 33:101566. <https://doi.org/10.1016/j.tmaid.2020.101566> PMID: 32007621
10. Hossain MM. Current Status of Global Research on Novel Coronavirus Disease (COVID-19): A Bibliometric Analysis and Knowledge Mapping. Available at SSRN: <https://ssrncom/abstract=3547824> or <http://dxdoiorg/102139/ssrn3547824> 2020.
11. Gianola S, Gasparini M, Agostini M, Castellini G, Corbetta D, Gozzer P, et al. Survey of the reporting characteristics of systematic reviews in rehabilitation. *Physical therapy.* 2013; 93(11):1456–66. <https://doi.org/10.2522/ptj.20120382> PMID: 23744458
12. Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. *PLoS medicine.* 2016; 13(5):e1002028. <https://doi.org/10.1371/journal.pmed.1002028> PMID: 27218655
13. Grant C, Williams B, Driscoll T. Historical trends in publications in the *International Journal of Epidemiology*. *International journal of epidemiology.* 2018; 47(3):938–41. <https://doi.org/10.1093/ije/dyy109> PMID: 29893853
14. Lin RT, Soeberg MJ, Chien LC, Fisher S, Takala J, Lemen R, et al. Bibliometric analysis of gaps in research on asbestos-related diseases: declining emphasis on public health over 26 years. *BMJ open.* 2018; 8(7):e022806. <https://doi.org/10.1136/bmjopen-2018-022806> PMID: 30049702
15. Muthuvel T, Isaakidis P, Shewade HD, Kattuppara L, Singh R, Govindarajulu S. Leprosy trends at a tertiary care hospital in Mumbai, India, from 2008 to 2015. *Global health action.* 2016; 9:32962. <https://doi.org/10.3402/gha.v9.32962> PMID: 27885973
16. Ligthelm RJ, Borzi V, Gumprecht J, Kawamori R, Wenying Y, Valensi P. Importance of observational studies in clinical practice. *Clinical therapeutics.* 2007; 29(6 Pt 1):1284–92. <https://doi.org/10.1016/j.clinthera.2007.07.004> PMID: 18036390
17. Alexander PE, Debono VB, Mammen MJ, Iorio A, Aryal K, Deng D, et al. COVID-19 coronavirus research has overall low methodological quality thus far: case in point for chloroquine/hydroxychloroquine. *Journal of clinical epidemiology.* 2020.
18. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews.* 2015; 4:1. <https://doi.org/10.1186/2046-4053-4-1> PMID: 25554246

19. Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. *Lancet*. 2011; 377(9760):108–9. [https://doi.org/10.1016/S0140-6736\(10\)60903-8](https://doi.org/10.1016/S0140-6736(10)60903-8) PMID: 20630580
20. MedRxiv. <https://www.medrxiv.org/>. accessed on June 4 2020.
21. Johansson MA, Reich NG, Meyers LA, Lipsitch M. Preprints: An underutilized mechanism to accelerate outbreak science. *PLoS medicine*. 2018; 15(4):e1002549.
22. Gutierrez J, Kuperman E, Kaboli PJ. Using Telehealth as a Tool for Rural Hospitals in the COVID-19 Pandemic Response. *The Journal of rural health: official journal of the American Rural Health Association and the National Rural Health Care Association*. 2020.
23. Kelly J, Sadeghieh T, Adeli K. Peer Review in Scientific Publications: Benefits, Critiques, & A Survival Guide. *Ejifcc*. 2014; 25(3):227–43. PMID: 27683470
24. Database PTWsMCB. <https://library.sacredheart.edu/c.php?g=29690>. accessed on June 4 2020.
25. Halladay CW TT, Schmid IT, Schmid CH, Dahabreh IJ. Using data sources beyond PubMed has a modest impact on the results of systematic reviews of therapeutic interventions. *Journal of clinical epidemiology*. 2015:1076–84. <https://doi.org/10.1016/j.jclinepi.2014.12.017> PMID: 26279401
26. Diana K. How preprint servers are blocking bad coronavirus research. *Nature*. 2020; 581:130–1. <https://doi.org/10.1038/d41586-020-01394-6> PMID: 32382120
27. Kwon D. How swamped preprint servers are blocking bad coronavirus research. *Nature* 2020; Vol 581
28. Registers WP. <https://www.who.int/ictrp/network/primary/en/>. accessed on June 4 2020.
29. Kambhampati SBS, Vaishya R, Vaish A. Unprecedented surge in publications related to COVID-19 in the first three months of pandemic: A bibliometric analytic report. *Journal of clinical orthopaedics and trauma*. 2020.
30. Gori D, Boetto E, Fantini M. The early scientific literature response to the novel Coronavirus outbreak: who published what? <https://doi.org/10.1101/2020032520043315>. 2020.
31. Fidahic M, Nujic D, Runjic R, Civljak M, Markotic F, Lovric Makaric Z, et al. Research methodology and characteristics of journal articles with original data, preprint articles and registered clinical trial protocols about COVID-19. *BMC Med Res Methodol*. 2020; 20(1):161. <https://doi.org/10.1186/s12874-020-01047-2> PMID: 32571302
32. Yang Z, Yongyue W, Sipeng S, Mingzhi Z, Feng C. Appealing for Efficient, Well Organized Clinical Trials on COVID-19. <https://doi.org/10.21037/atm-20-2429> PMID: 32566569
33. Luo Q, Qin T. Managing clinical trials for covid-19: the importance of ethics committees. *Bmj*. 2020; 369:m1369. <https://doi.org/10.1136/bmj.m1369> PMID: 32245769
34. Khachfe H, Akl E. A large number of COVID-19 interventional clinical trials were registered soon after the pandemic onset: a descriptive analysis. *Journal of clinical epidemiology*. 2020.
35. Gao J, Tian Z, Yang X. Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci Trends*. 2020; 14(1):72–3. <https://doi.org/10.5582/bst.2020.01047> PMID: 32074550
36. Ektorp E. Death threats after a trial on chloroquine for COVID-19. *Lancet Infect Dis*. 2020; 20(6):661.
37. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *Journal of clinical epidemiology*. 2011; 64(12):1303–10. <https://doi.org/10.1016/j.jclinepi.2011.04.014> PMID: 21802903
38. Hodgson J. The pandemic pipeline. *Nature biotechnology*. 2020; 38(5):523–32. <https://doi.org/10.1038/d41587-020-00005-z> PMID: 32203293
39. Ebuenyi ID, Syurina EV, Bunders JFG, Regeer BJ. Barriers to and facilitators of employment for people with psychiatric disabilities in Africa: a scoping review. *Global health action*. 2018; 11(1):1463658. <https://doi.org/10.1080/16549716.2018.1463658> PMID: 29764342
40. Graham F. Daily briefing: medRxiv brings preprints to medical science. *Nature Briefing*. 2019.
41. Vabret N, Samstein R, Fernandez N, Merad M, Sinai Immunology Review P, Trainees, et al. Advancing scientific knowledge in times of pandemics. *Nature reviews Immunology*. 2020; 20(6):338. <https://doi.org/10.1038/s41577-020-0319-0> PMID: 32327718
42. Wpro. Coronavirus disease (COVID-19) Call for Papers. *Western Pacific Surveillance and Response*. <https://ojs.wpro.who.int/ojs/index.php/wpsar>.
43. Iwema CL, LaDue J, Zack A, Chattopadhyay A. search.bioPreprint: a discovery tool for cutting edge, preprint biomedical research articles. *F1000Research*. 2016; 5:1396. <https://doi.org/10.12688/f1000research.8798.2> PMID: 27508060
44. Glasziou P, Chalmers I. Comment: Can it really be true that 50% of research is unpublished? accessed on June 4 2020. 2017; <https://blogs.bmj.com/bmj/2017/06/05/paul-glasziou-and-iain-chalmers-can-it-really-be-true-that-50-of-research-is-unpublished/>.

45. Waste LssioR. <https://www.thelancet.com/series/research>. accessed on June 4 2020.
46. <https://en.wikipedia.org/wiki/Preprint>.
47. Lariviere V, Sugimoto CR, Macaluso B, Milojevic S, Cronin B, Thelwall M. arXiv e-prints and the journal of record: An analysis of roles and relationships. accessed on June 4 2020. 2020.
48. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of autoimmunity*. 2020; 109:102433. <https://doi.org/10.1016/j.jaut.2020.102433> PMID: [32113704](https://pubmed.ncbi.nlm.nih.gov/32113704/)
49. Fu DY, Hughey JJ. Releasing a preprint is associated with more attention and citations for the peer-reviewed article. *eLife*. 2019; 8.
50. Lv M, Luo X, Estill J, Liu Y, Ren M, Wang J, et al. Coronavirus disease (COVID-19): a scoping review. *Euro surveillance: bulletin European sur les maladies transmissibles = European communicable disease bulletin*. 2020; 25(15).
51. Plos. <https://plos.org/open-science/preprints/>. accessed on June 4 2020.
52. Rzymiski P, Nowicki M, Mullin GE, Abraham A, Rodriguez-Roman E, Petzold MB, et al. Quantity does not equal quality: Scientific principles cannot be sacrificed. *Int Immunopharmacol*. 2020; 86:106711. <https://doi.org/10.1016/j.intimp.2020.106711> PMID: [32585609](https://pubmed.ncbi.nlm.nih.gov/32585609/)
53. Charles P, Kelly S. Two elite medical journals retract coronavirus papers over data integrity questions. *Science*. 2020 June 4.
54. f1000research.com. search.bio. Preprint: a discovery tool for cutting edge, preprint biomedical research articles [version 2; peer review: 2 approved]. <https://f1000research.com/articles/5-1396/v2>. accessed on June 12 2020.
55. Bauchner H, Fontanarosa PB, Golub RM. Editorial Evaluation and Peer Review During a Pandemic: How Journals Maintain Standards. *JAMA*. 2020.
56. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Obstet Gynecol*. 2009; 114(6):1341–5. <https://doi.org/10.1097/AOG.0b013e3181c3020d> PMID: [19935040](https://pubmed.ncbi.nlm.nih.gov/19935040/)