Taylor & Francis
Taylor & Francis Group

Accountability in Research
Policies and Quality Assurance

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/gacr20

An alarming retraction rate for scientific publications on Coronavirus Disease 2019 (COVID-19)

Nicole Shu Ling Yeo-Teh & Bor Luen Tang

To cite this article: Nicole Shu Ling Yeo-Teh & Bor Luen Tang (2021) An alarming retraction rate for scientific publications on Coronavirus Disease 2019 (COVID-19), Accountability in Research, 28:1, 47-53, DOI: 10.1080/08989621.2020.1782203

To link to this article: https://doi.org/10.1080/08989621.2020.1782203

ACCOUNTABILITY IN RESEARCH
2021, VOL. 28, NO. 1, 47–53
https://doi.org/10.1080/08989621.2020.1782203

Taylor & Francis
Taylor & Francis Group

COMMENTARY

An alarming retraction rate for scientific publications on Coronavirus Disease 2019 (COVID-19)

Nicole Shu Ling Yeo-Teh [a] and Bor Luen Tang [a,b]

[a]Research Compliance and Integrity Office, National University of Singapore, Singapore, Republic of Singapore; [b]NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore, Republic of Singapore

ABSTRACT

The ongoing Coronavirus Disease 2019 (COVID-19) global pandemic has triggered a flurry of associated research publications, numbering to ~137 papers a day since February 2020. This rate of publication appears to be exceptionally high, when compared to research papers published on other similar topics. Searches of COVID-19-associated publications on PubMed and Retraction Watch Database indicate that the retraction record appearance rate for COVID-19-related research is also exceptionally high compared to other related research topics in viral epidemics/pandemics and surpasses the basal level of about 4 in 10,000 papers. This finding serves as a reminder and caution against any lapses in the standard of work, peer review, and publication of COVID-19-related research.

KEYWORDS

Coronavirus Disease 2019 (COVID-19); pandemic; research ethics; retractions

Introduction

The ongoing Coronavirus Disease 2019 (COVID-19) global pandemic (Kakodkar, Kaka, and Baig 2020) is the most widespread and most severe in modern times. The response of the scientific community, evidenced by the large number of research publications, has been unprecedentedly swift. Since the World Health Organization (WHO) declared it a "Public Health Emergency of International Concern" on 30 January 2020, and recognized it as a pandemic on 11 March 2020, thousands of papers on various aspects of the disease and its causative agent, the Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), have appeared in literature. This very high rate of publication on COVID-19 research is undoubtedly aided by the call for data sharing (Wellcome Trust 2020) and with most major publishers collating and allowing open access to all papers on the topic. COVID-19 research papers are also afforded expedited processing in some journals such as

CONTACT Nicole Shu Ling Yeo-Teh dprysln@nus.edu.sg; Bor Luen Tang dprtbl@nus.edu.sg Research Compliance and Integrity Office, National University of Singapore, Singapore, Republic of Singapore

Personality and Individual Differences (PAID), an Elsevier journal,[1] and manuscripts submitted to Hindawi journals enjoy article processing fee waiver.[2] These actions by journals are also congruent with the goals stated in WHO's R&D Blueprint,[3] a roadmap to accelerate research concerning this viral pandemic.

A comparison of publications stemming from COVID-19 to those from other pandemics

A search of National Center for Biotechnology Information (NCBI)'s PubMed database (https://pubmed.ncbi.nlm.nih.gov/) with the term `COVID-19′ in the title and/or abstracts retrieved a list of 7440 articles (as of 3 May 2020) and 17,559 articles (as of 8 June 2020). A closer scrutiny of the earlier articles indicates that the earliest COVID-19 papers on PubMed appeared in February 2020. A similar search with the term `H1N1′ from the period of June 2009 to May 2010, an entire year of the Influenza A H1N1 pandemic (pandemic declared by WHO in June 2009) (Richt, Webby, and Kahn 2013), retrieved only 2316 papers. To date, the average rate by which papers on COVID-19 appears in PubMed was therefore about 137 papers per day, compared to the average of 6 papers per day on H1N1 in the year following its declaration as a pandemic. By June 2020, COVID-19 has only been identified to be a serious health problem for just under 6 months. This is barely more than the 5.5 months which some journals claim it would take from first submission to online publication. Yet, over 17,000 publications were fortunate enough to have such expedient publication processes.

The high rate of COVID-19 publications attests to the urgent and intense effort by the scientific community to better understand the virus, its transmission, and COVID-19 disease pathology; and we have indeed learned much (Hirano and Murakami 2020; Li, Liu, and Ge 2020). However, there is a concern that such a high rate and intensity of additions to the literature may well be associated with inadvertent errors and inaccuracies, as well as varying degrees of misinformation.

We searched the Retraction Watch database (RWD) (http://retractiondatabase.org/RetractionSearch.aspx) for retraction records associated with COVID-19 publications. Using the keyword `COVID-19′ in the `Title' search section retrieved six items (on 3 May 2020) and 17 items (on 8 June 2020), which corresponded to 5 and 13 unique articles on record, respectively. These have either been fully retracted/withdrawn, tagged with editorial expression of concerns or issued with a corrigendum. In one of the retraction notices available online, the authors specifically "… deeply apologize for this premature publication."[4] Other retraction notices cite reasons of conclusions being reached from theoretical deductions rather than empirical evidence, and the recount of Chinese medical staff's experiences from a secondary

source. Another retraction notice worryingly states that the authors "lack of better explanations of the inclusion criteria and the triage of patients to ensure patient safety." It is worth noting that top journals are not spared, since articles concerned include those from prominent medical journals such as *Lancet* and *New England Journal of Medicine*.

The retraction record (inclusive of retractions/withdrawals, corrigendums, and expression of concerns), based on the total number of papers retrieved using the keyword `COVID-19′ from PubMed, is thus 0.074%. Notably, papers on a COVID-19 therapy based on an anti-malarial drug chloroquine/hydroxychloroquine (in combination or otherwise with an antibiotic azithromycin) have received much publicity due to its anecdotal promotion by influential figures.

To see how this COVID-19-related retraction rate compares with publications in related research fields, we performed similar PubMed and RWD searches with other keywords, including `MERS', `H1N1′, `HIV,' `Ebola,' and `Virus' (searched under "Title" of RWD). To gauge the general retraction record in the broader field, we also did a basic PubMed search with the general terms `Immunology' and `Cancer,' and retractions under "Subject" of RWD. The results for these searches (as per 8 June 2020) are tabulated in Table 1. The calculated retraction record appearance rate for titles bearing the different virus epidemic/pandemic keywords ranges from 0.021% to 0.024%. For articles included under the subject of `Immunology' and `Cancer' in RWD, the retraction rates are 0.037% and 0.039%, respectively.

Table 1. A comparison of retraction rates between papers on COVID-19 and those of related and general fields (as per 8 June 2020). For COVID-19, an earlier search done on 3 May 2020[a] is shown for comparison. PubMed searches were conducted by limiting the keyword in the Title/Abstract field, except for 'Cancer' and 'Immunology', which were carried out as a plain search. RWD searches were carried out with the keyword in the Title, except for 'Cancer' and 'Immunology' searches which were done under the field of "Subject".

Keywords used for search	PubMed	Items (unique) in RWD	Retraction rate (%, 3 d.p.)
COVID-19 (as of 3 May 2020)[a]	7,440	6 (5)	**0.081 (0.067)**
COVID-19 (as of 8 June 2020)	17,559	17 (13)	**0.097 (0.074)**
Ebola	8,428	2 (2)	0.024
H1N1	17,378	4 (4)	0.023
HIV	311,888	72 (64)	0.023 (0.021)
MERS[b]	4,121	1 (1)	0.024
Virus	671,832	164 (153)	0.024 (0.023)
Cancer	4,119,529	1,616	0.039
Immunology	1,773,526	642	0.037

[a]Items retrieved by searching RWD under "Title" are manually curated for duplicates, and the unique number of publications are shown in brackets.

It would be obvious to include `Severe acute respiratory syndrome (SARS)'. However, in an RWD search for retractions, the full disease name retrieved no items while its abbreviation `SARS' retrieved 5 items, all of which pertain to SARS-CoV-2, the virus of COVID-19.

[b]For MERS, we used a search term `MERS [Title/Abstract] BUTNOT k-mers [Title/Abstract] BUTNOT n-mers [Title/Abstract]' that would exclude the terms `n-mers' and `k-mers' commonly used in unrelated bioinformatics papers.

These latter numbers are in close agreement with a previous estimate of about four retractions for every 10,000 papers (~0.04%) (Brainard and You 2018). It is apparent that the retraction record appearance rate for COVID-19-related research to date is exceptionally high.

Why do COVID-19-associated publications have such a high retraction record appearance?

There may be several technical reasons for the high retraction record appearance of COVID-19-related publications. The first is a sampling issue, given that the COVID-19 pandemic is still ongoing/evolving and related papers are still being produced at a high rate. The second is that there may be a significant discrepancy in coverage congruence between the papers retrieved using the same keyword in PubMed and RWD. For the first possibility, it should be noted that the number of publications associated with COVID-19 has now surpassed those for Ebola, H1N1 and MERS, and sampling is likely not an issue. Despite the large number of publications, the number of retracted papers is still seemingly high. Pertaining to the second possibility, we do not think that keyword search coverage congruence is an issue that would stand out specifically for COVID-19 since the database of our choice indexes most biomedical journals, in which COVID-19-related articles would be included.

Discounting the above possible technical issues, we are then left with having to account for the high retraction rate of COVID-19-related research. Notably, the numbers obtained are based only on published journal papers and a small number of preprints. There are other forms of publication, such as a genome sequence submission that has greatly exaggerated the viral mutation rate,[5] as well as mistaken claims of similarity between the SARS-CoV-2 spike protein to HIV-1 gp120 and Gag in a preprint,[6] that have since been retracted. Given that retraction of journal articles typically takes time (Trikalinos, Evangelou, and Ioannidis 2008; Dal-Ré and Ayuso 2019), this unusually high rate of retraction of COVID-19-related articles, with an average retraction time of less than 2 weeks, occurring in a matter of 4 months may well be a significant underestimate of the greater numbers in the pipeline.

Responsibility and accountability of researchers, authors and journals

If so, what is happening? Could there have been lapses in stringency and standards amid the rush to produce results and to publish? Could the sense of urgency to share any findings, no matter how preliminary, that strive to curb the pandemic, overwhelm the typically careful and reserved nature of

scientific research review and dissemination? Or could authors be exploiting the thirst for articles relating to COVID-19 of any kind? All the above may have played a role in the apparently high retraction record appearance rate of COVID-19 publications.

Research into COVID-19 diagnostics and therapeutics has been given top priority in many affected countries with any capacity to conduct epidemiological, clinical, and basic sciences research associated with the disease and its causative agent. Such COVID-19 research has continued despite lock-down measures in many countries that affect most other research activities, with public grant funding acutely re-focused on COVID-19. The call to arms to tackle the disease and the virus creates the pressure to justify support and generates hyper competition among researchers. It is also a possibility that enticed by funding availability, some researchers could have switched fields to work on the "hot topic" of the generation. On the other hand, with the high flux of submissions, the peer review system is in turn stressed. The limited number of truly qualified peer reviewers is themselves working hard on their own research and papers, which may result in the task of reviewing COVID-19 manuscripts being taken on by reviewers who are not particularly equipped with domain expertise, and therefore not as perceptively critical. Furthermore, the proliferation of preprint submissions has greatly enhanced the visibility of papers which have not undergone peer review. While there are no data yet to indicate that a pre-exposure at preprint level enhances acceptance of a manuscript in a journal, one of us (BLT) has a personally seen authors using the `views and downloads' number on a preprint to argue for a manuscript's acceptance.

The numbers we observe here would serve as a reminder and caution against any lapses in the standard of work, peer review, and publication of COVID-19-related research. Accuracy and trustworthiness of published results are particularly important in a global emergency like such. These results may incur common beliefs in the larger population, particularly patients and health-care workers who are desperate for any form of therapeutics, especially those that can be used in an off-label manner. Take the reported treatment with chloroquine/hydroxychloroquine for example, its efficacy and safety was unproven (Fihn, Perencevich, and Bradley 2020). A Brazilian trial on the treatment has since been halted due to patient death (Reuters Fact Check team 2020). An editorial expression of concern has been issued on a *Lancet* paper on such a treatment strategy (The Lancet Editors 2020), and the paper has since been retracted. A recent paper has indicated that there is no evidence for rapid antiviral clearance or clinical benefit with the combination of hydroxychloroquine and azithromycin in COVID-19 patients (Molina et al. 2020). A preprint posted on a retrospective analysis of data from COVID-19 patients hospitalized in all United States Veterans Health Administration medical centers until 11 April 2020 "… found no evidence that use of hydroxychloroquine, either with or without azithromycin, reduced the risk of mechanical ventilation in patients hospitalized with COVID-19." In fact, the authors found "… an association of increased overall mortality was identified in patients treated with hydroxychloroquine alone" (Magagnoli et al. 2020). However, public

belief in the treatment has already claimed casualties, and may have also affected enrollment in clinical trials for other potential therapies (Ledford 2020). In should be borne in mind that therapeutic interventions and approaches informed by poor data and inaccurate results could critically mislead and cost the already stressed health systems and patients dearly. Furthermore, effort and resources spent on following up from unreliable findings on the biology of COVID-19 would be rendered futile and a waste of resources. Ensuring accuracy and rigor of publications associated with COVID-19, as with all other research, is a collective responsibility of researchers, would-be authors, as well as the journals.

Conclusion

We observed an alarmingly high rate of retraction associated with COVID-19 publications, which may be due to a rush by researchers to submit papers for publications and the ease by which some of these are expedited for publication. In a global disaster like the COVID-19 pandemic, where human resilience and recovery are crucially dependent on solutions that would be provided by science and research, it is even more important for the latter to be conducted with the utmost stringency and integrity. Publishing of research works in this regard should proceed with the same rigor and this is the collective responsibility of researchers and publishers alike.

Notes

1. https://www.journals.elsevier.com/personality-and-individual-differences/news/priority-handling-for-covid-19-papers
2. https://www.hindawi.com/covid-19-collection/
3. https://www.who.int/who-documents-detail/a-coordinated-global-research-roadmap
4. https://www.bdd.rdplf.org/index.php/bdd/article/view/54713
5. Sequence from virus collected from a female COVID-19 patient in St Petersburg and uploaded to the GISAID database (https://www.gisaid.org/) carried 200 mutations compared with the first sample released by Chinese researchers in early January. This was soon retracted and updated with one that indicated 7 mutations.
6. Pradhan et al. Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag. BioR×iv. doi: https://doi.org/10.1101/2020.01.30.927871 (withdrawn)

Disclosure statement

No conflict of interest.

ORCID

Nicole Shu Ling Yeo-Teh http://orcid.org/0000-0002-9153-9663
Bor Luen Tang http://orcid.org/0000-0002-1925-636X

References

Brainard, J., and J. You. 2018. "What a Massive Database of Retracted Papers Reveals about Science Publishing's 'Death Penalty'." https://www.sciencemag.org/news/2018/10/what-massive-database-retracted-papers-reveals-about-science-publishing-s-death-penalty

Dal-Ré, R., and C. Ayuso. 2019. "Reasons for and Time to Retraction of Genetics Articles Published between 1970 and 2018." *Journal of medical genetics* 56: 734–740. doi:10.1136/jmedgenet-2019-106137.

Fihn, S. D., E. Perencevich, and S. M. Bradley. 2020. "Caution Needed on the Use of Chloroquine and Hydroxychloroquine for Coronavirus Disease 2019." *JAMA Network Open* 3 (4): e209035. doi:10.1001/jamanetworkopen.2020.9035.

Hirano, T., and M. Murakami. 2020. "COVID-19: A New Virus, but A Familiar Receptor and Cytokine Release Syndrome." *Immunity* 52 (5): 731–733. doi:10.1016/j.immuni.2020.04.003.

Kakodkar, P., N. Kaka, and M. N. Baig. 2020. "A Comprehensive Literature Review on the Clinical Presentation, and Management of the Pandemic Coronavirus Disease 2019 (COVID-19)." *Cureus* 12: e7560. doi:10.7759/cureus.7560.

The Lancet Editors. 2020. "Expression of Concern: Hydroxychloroquine or Chloroquine with or without A Macrolide for Treatment of COVID-19: A Multinational Registry Analysis." *Lancet* S0140-6736 (20): 31290–31293. doi:10.1016/S0140-6736(20)31290-3.

Ledford, H. 2020. "Chloroquine Hype Is Derailing the Search for Coronavirus Treatments." https://www.nature.com/articles/d41586-020-01165-3

Li, H., Z. Liu, and J. Ge. 2020. "Scientific Research Progress of COVID-19/SARS-CoV-2 in the First Five Months." *Journal of Cellular and Molecular Medicine*. doi:10.1111/jcmm.15364.

Magagnoli, J., S. Narendran, et al. 2020. "Outcomes of Hydroxychloroquine Usage in United States Veterans Hospitalized with Covid-19." *medRxiv*. doi:10.1101/2020.04.16.20065920v2.

Molina, J. M., C. Delaugerre, J. Le Goff, B. Mela-Lima, D. Ponscarme, L. Goldwirt, and N. de Castro. 2020. "No Evidence of Rapid Antiviral Clearance or Clinical Benefit with the Combination of Hydroxychloroquine and Azithromycin in Patients with Severe COVID-19 Infection." *Médecine et Maladies Infectieuses* 50 (4): 384. doi:10.1016/j.medmal.2020.03.006.

Reuters Fact Check team. 2020. "True Claim: Brazilian Chloroquine Study on COVID-19 Patients Was Halted after 11 Deaths." https://www.reuters.com/article/uk-factcheck-brazilian-chloroquine-study/true-claimbrazilian-chloroquine-studyon-covid-19-patients-washalted-after11deaths-idUSKBN21Z395

Richt, J. A., R. J. Webby, and R. E. Kahn. 2013. "The Pandemic H1N1 Influenza Experience." *Current Topics in Microbiology and Immunology* 365: 269–279. doi:10.1007/82_2013_309.

Trikalinos, N. A., E. Evangelou, and J. P. A. Ioannidis. 2008. "Falsified Papers in High-impact Journals Were Slow to Retract and Indistinguishable from Nonfraudulent Papers." *Journal of Clinical Epidemiology* 61 (5): 464–470. doi:10.1016/j.jclinepi.2007.11.019.

Wellcome Trust. 2020. "Sharing Research Data and Findings Relevant to the Novel Coronavirus (COVID-19) Outbreak." https://wellcome.ac.uk/coronavirus-covid-19/open-data