Bibliometric analysis of coronavirus disease (COVID-19) literature published in Web of Science 2019–2020

Rai K. Farooq, Shafiq Ur Rehman¹, Murtaza Ashiq², Nadeem Siddique³, Shakil Ahmad¹

Abstract:
Coronavirus outbreak in Wuhan, China, turned into a pandemic in record time. Communication of disease presentation and mechanism of spread remain keys to getting ahead of the virus and limiting its spread beyond the capacity of management. Owing to huge academic focus and pandemic concern around the globe, this bibliometric analysis investigated research productivity related to coronavirus disease (COVID-19) pandemic using the Web of Science database. The relevant data were harvested, and search query was further refined by publication years (2020 OR 2019) and document types (article, book chapter, and proceedings paper). Finally, 6694 records were imported and downloaded in Plaintext and BibTeX formats on August 1, 2020. The data analysis was performed using MS Excel, VOS viewer, and Biblioshiny software. The pattern of multi-author publications has outstripped that of single-authors. Apart from COVID-19 and the novel coronavirus, the important keywords mentioned included pandemic, pneumonia, epidemiology, public health, outbreak, epidemic, China, infection, and treatment. The analysis shows a strong local research response from China, with large teams reporting on the disease outbreak. Subsequent studies will document a global response as the virus spreads worldwide. The initial research related to the current coronavirus outbreak was reported from within China. The data and patterns were supposed to alter as the virus spread globally.

Keywords:
2019-nCoV, bibliometric, bibliometric-coronavirus, COVID-19, health care, pandemic, China, severe acute respiratory syndrome-cov-2, research productivity, the World Health Organization

Introduction

In today’s connected world, contagious and previously unknown pathogens have become a public health problem. In late December 2019, the health facilities of Wuhan city in China reported receiving patients with chest infection of an unknown cause. The clinical picture of these patients resembled viral chest infections and included fever, cough, and dyspnea. Initial laboratory investigations linked it with the Coronavirus family; however, no confirmatory links were established.¹ The World Health Organization (WHO) was alarmed about this outbreak. Within a month (January 30, 2020), the WHO had declared this outbreak a global health emergency. On March 11, 2020, they pronounced it a pandemic.² In the absence of a specific treatment or vaccine, minimizing human contact was the only way to slow its spread.³ The virus has been named a novel coronavirus and coded as COVID‑19 and 2019‑nCoV. Phylogenetic analyses have revealed its close similarity with the severe acute respiratory syndrome-cov-2, research productivity, the World Health Organization

How to cite this article: Farooq RK, Rehman SU, Ashiq M, Siddique N, Ahmad S. Bibliometric analysis of coronavirus disease (COVID-19) literature published in Web of Science 2019–2020. J Fam Community Med 2021;28:1-7.
Farooq, et al.: Bibliometric analysis of COVID‑19 pandemic, bibliometric data were retrieved from the Web of Science (WOS), the most authentic indexing and abstracting database in the world. In the WOS core collection, the following search query was run in title and Author Keyword field.

TI= (coronavirus OR covid19 OR covid‑19 OR ncov2019 OR sars‑cov‑2 OR “SARS COV 2″ OR orthocoronavirinae) OR AK = (coronavirus OR covid19 OR covid‑19 OR ncov2019 OR sars‑cov‑2 OR “SARS COV 2″ OR orthocoronavirinae).

The query was further refined by publication years (2020 OR 2019) and document types (article, or review article or book chapter or proceedings paper or early access). Letters, data paper, news items, and editorials were excluded. Finally, the 6694 records (articles = 3917, early access articles = 1487, review articles = 1023 review, early access review articles = 263, book chapters = 3, and proceeding paper = 1) were imported and downloaded in Plaintext and BibTeX formats on July 9–August 1, 2020. The citation impact (CI) used in this study was calculated by dividing the total number of citations by the total number of publications. This illustrates the average number of citations that a specific publication has received. Besides, some terminologies are defined by the WOS database, such as usage count U1 which refers to the usage count of the last 180 days. The data analysis was performed using MS Excel, MS Access, and various bibliometric softwares, including VOS viewer, Biblioshiny (RStudio), and BibExcel.

**Results**

In 2019, 324 publications obtained 2105 citations and a usage count of 2776 whereas from January to July 2020, i.e., 180 days, 6370 publications were recorded with 37,965 citations and a usage count of 46,885. This shows that the COVID‑19 pandemic has had remarkable attention from researchers throughout the world.

**Most Productive Countries and Organizations**

Table 1 indicates top 10 countries and organizations that produced COVID‑19 literature globally. All countries in the table produced over 100 publications from 2019 to July 1, 2020. Only four countries produced over 500 publications. The United States of America (USA) is at the top of the list with 1860 publications, 9468 citations followed by China with 1510, Italy with 782, and England with 592 publications. It has been observed that though the USA has a higher number of publications, the impact of publications by China is higher than any country. Of the top 10 organizations, Huazhong University Science and Technology produced 193 publications and 5484 citations, followed by Wuhan University with 133...
publications and 5752 citations. The Chinese Academy of Sciences, China, ranks 7th with 82 publications.

### Influential Research Journals

Table 2 presents the top 10 research journals that produce literature on COVID-19. Nine journals produced over 50 publications; two of those journals produced over 100 publications. Journal of Medical Virolology (Quartile 4) emerged as a top source with 149 publications, 1631 citations, followed by CUREUS (Cureus Journal of Medical Science), a nonimpact factor journal with 142 publications and 92 citations. The source with the highest impact factor (6.551) “Science of the Total Environment” produced 77 publications and obtained 170 citations. Most of the journals belonged to the USA (3), the Netherlands (3), and Switzerland (2), while India, and France had one journal each.

### Authorship Pattern

The analysis of authorship pattern shows that all top ten authorship patterns have a significant number of publications. The top three authorship patterns were two authors (830 publications), three authors (763 publications), and four authors (718 publications). It is noteworthy that a significant number of publications (712) have single authorship. There was a decline in number of studies for more than 5 authors with 519, 474, 370, 268, and 228 publications recorded for 6, 7, 8, 9, and 10 authors, respectively.

### Authors’ Keyword Analysis

Figure 1 presents the authors’ keyword analysis on COVID-19 literature from VOS viewer software. The minimum occurrence of 30 was selected; hence, 75 keywords met this criterion consisting of seven clusters. The size of the bubble indicates the number of occurrence and total strength links with other items/keywords. The top five keywords were COVID-19, coronavirus, sars-cov-2, pandemic, and pneumonia, which occurred 3409, 1201, 1266, 492, and 189 times, respectively.

### Highly Cited Articles

Table 3 highlights the top ten highly cited articles on COVID-19. It is interesting that all the top ten articles were published in 2020. Half of those articles were published in Lancet, three in The New England Journal of Medicine, and one in The Journal of American Medical Association and Nature each. Half of the articles got over 1000 citations. The article entitled “Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China” by “Huang CL” published in Lancet got the highest citations (2264), U1 score (997), Z9 score (2396), and TR (37). The article at the bottom of the list by ‘Lu RJ’ received 639 citations.

### Three-Factor Analyses of Major Aspects of the Data

**Countries, keywords, and organizations**

Figure 2 presents the three-factor analysis of the relationship among countries (left), keywords (middle), and authors’ affiliated organizations (right). It shows that six countries (China, USA, Italy, United Kingdom, India, France, and Germany) published COVID-19 literature mostly using four main keywords (COVID-19, coronavirus, sars-cov-19, pandemic). These countries and keywords have a strong relationship with five organizations (Huazhong university science and technology China, Wuhan University China, Fudan University China, Tehran University of Medical Sciences, and the University of Hong Kong).

**Country Collaboration Map on COVID-19 Literature**

Figure 3 presents the country collaboration map on COVID-19 literature around the world. There are 1817 entries of collaborations among various countries.
worldwide with a maximum of 272 to one collaboration. The United States of America and China are top collaborating countries with 272 collaborations, followed by the USA and UK with 140 collaborations, Italy and the USA with 136 collaborations, the USA and Canada with 133 collaborations, and Italy and the UK with 130 collaborations.

Discussion

Bibliometric analysis is increasingly being used for the review of trends and progress in different fields and areas of research. The current analysis of data represents different dimensions of COVID-19 research, which includes the top countries, organizations, and journals producing publications on COVID-19. Of these countries, the USA stands as the number one country in terms of research on COVID-19 closely followed by China and distantly by such countries as Italy, England, and India. Although China has produced relatively fewer publications than the top producing country, its publications have obtained significantly higher CI than any other country listed. It is also notable that despite being ranked 5th in terms of the number of publications, India is ranked outside the top 10 countries with a CI 2.03. This puts a huge responsibility on Indian policymakers and scientists to fund more meaningful and citable research. Our results are in line with the findings of Tao, Zhou [12] who reported the USA as the most productive country. This recent trend outstrips the statistics noted by

Table 2: Top ten most highly influential research journals on COVID-19 literature all over the world during 2019–2020

| Source                                      | TP  | TC   | H_index | Impact factor | Quartile | Publisher       | Country     |
|---------------------------------------------|-----|------|---------|---------------|----------|-----------------|-------------|
| Journal of Medical Virology                | 149 | 1631 | 24      | 2.021         | 4        | Wiley-Blackwell  | USA         |
| Cureus (Cureus Journal of Medical Science) | 142 | 92   | 5       | N/A           | N/A      | Cureus          | USA         |
| Head and Neck                              | 85  | 135  | 5       | 2.538         | 1        | Wiley           | USA         |
| Science of the Total Environment           | 77  | 170  | 7       | 6.551         | 1        | Elsevier        | Netherlands |
| Journal of Clinical Virology               | 62  | 63   | 4       | 2.777         | 3        | Elsevier        | Netherlands |
| Eurosurveillance                           | 60  | 641  | 10      | 6.454         | 1        | ECDC            | France      |
| International Journal of Infectious Diseases | 54  | 563  | 12      | 3.202         | 2        | Elsevier        | Netherlands |
| International Journal of Environmental Research and Public Health | 53  | 176  | 6       | 2.468         | 2        | MDPI            | Switzerland |
| Viruses-Basel                               | 52  | 520  | 13      | 3.816         | 2        | MDPI            | Switzerland |
| Journal of Pure and Applied Microbiology   | 48  | 10   | 3       | N/A           | N/A      | Journal of Pure and Applied Microbiology | India |

NA=Not available, ECDC=European Centre for Disease Prevention and Control, MDPI=Multidisciplinary Digital Publishing Institute, TP=Total Publications TC=Total Citations

Table 3: Top ten most highly cited articles by researchers in the world during 2019–2020

| Title                                                                 | Author            | Year | Source                                      | TC  | U1  | TC per year |
|-----------------------------------------------------------------------|-------------------|------|---------------------------------------------|-----|-----|-------------|
| Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China | Huang Cl          | 2020 | Lancet                                      | 2264| 997 | 2264        |
| Clinical characteristics of 138 hospitalized Patients with 2019 novel coronavirus-infected Pneumonia in Wuhan, China | Wang DW           | 2020 | JAMA: The Journal of the American Medical Association | 1343| 230 | 1343        |
| Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus Pneumonia in Wuhan, China: A descriptive study | Chen NS           | 2020 | Lancet                                      | 1234| 392 | 1234        |
| Clinical characteristics of coronavirus disease 2019 In China          | Guan W            | 2020 | The New England Journal of Medicine         | 1226| 50  | 1226        |
| A novel coronavirus from patients with pneumonia in China, 2019       | Zhu N             | 2020 | The New England Journal of Medicine         | 1198| 631 | 1198        |
| A pneumonia outbreak associated with a new coronavirus of probable bat origin | Zhou P           | 2020 | Nature                                      | 871 | 12  | 871         |
| Early transmission dynamics in Wuhan, China, of Novel coronavirus-infected pneumonia | Li Q              | 2020 | The New England Journal of Medicine         | 841 | 34  | 841         |
| Clinical course and risk factors for mortality of adult inpatients with covid-19 in Wuhan, China: A retrospective cohort study | Zhou F            | 2020 | Lancet                                      | 790 | 281 | 790         |
| A familial cluster of Pneumonia associated with the 2019 Novel coronavirus indicating person-to-person transmission: A study of a family cluster | Chan JFW           | 2020 | Lancet                                      | 744 | 428 | 744         |
| Genomic characterization and epidemiology of 2019 Novel coronavirus: Implications for virus origins and receptor binding | Lu RJ             | 2020 | Lancet                                      | 639 | 412 | 639         |

TP=Total Publications TC=Total Citations
Bonilla-Aldana, Quintero-Rada[13] who used the Science Citation Index (SCI), Scopus, and PubMed databases with the term “Coronavirus” as the main operator from January 1951–January 2020. The study identified 18,158 articles from Scopus (31.3% from the USA, China 13.6%, and the United Kingdom 7.4%) followed by PubMed with 14,455 (20.1% the USA, China 18.6%, and Germany 4.2%), and SCI with 11,775 articles (34.9% from the USA, 22.4% China, and 6.8% Germany). This study also contradicts the findings of Chahrour et al.,[14] Hamidah et al.,[15] and Dehghanbanadaki et al.,[16] who ranked China as the top country that has produced research on COVID-19. This may be because COVID-19 was initially a public health problem in China.

Keeping in view the country-wise contribution on the subject, it is not surprising that out of the top 10 organizations listed, eight including the top two belong to China, followed by the USA with one organization.

Although Chinese Academy of Science produced a smaller number of publications, it secured the highest CI than all other organizations. Dehghanbanadaki et al.,[16] have ranked the University of Hong Kong and Huazhong University of Science and Technology the first in producing documents on the topic. This study contradicts the findings of Hossain,[17] who reported China as the top country and the University of Hong Kong as a top organization. This trend of CI indicates that people around the world are reading the original research produced by the organizations working at the epicenter of the disease. This also highlights the epidemiological/public health significance of sharing the data in addition to treating patients. Instant access to news and research has created a solid foundation for epidemiological debate, presenting some idea of the size of the problem, and encouraging preventive measures.

The bibliometric data regarding the most productive journal publishing research on COVID-19 discloses some interesting results. The Journal of Medical Virology (Wiley-Blackwell) has the most publications closely followed by CUREUS (Cureus Journal of Medical Science), which is the publication of Cureus, USA, and “Head and Neck” being published by Wiley, USA. The Dehghanbanadaki et al.,[16] reported that The Lancet and BMJ Clinical Research Ed were the most prolific in publishing documents on the topic. Lou et al.[18] conducted a PubMed based bibliometric analysis of COVID-19 literature and found that the Journal of Medical Virology was the most productive journal on the topic. These results also contradict the findings of Hossain,[17] who reported Viruses-Basel, Journal of Virology, and Transboundary and Emerging Diseases to be in the first, second, and third positions,
respectively in terms of their productivity. The Journal of Medical Virology has also attained the highest citations of 1631 while of the rest, only three journals Eurosurveillance, International Journal of Infectious Diseases, and Viruses-Basel, got over 500 citations. However, though CUREUS is in the second position on the list it has received only 92 citations so far. It is encouraging that the most relevant and authentic publication sources are getting more attention and referral. It is also evident that initial articles were prepared entirely locally with international attention and collaboration arriving later.

The analysis of data regarding authorship and collaborative research patterns shows that most publications are by two-author collaboration, closely followed by three, four, and single authors. Interestingly, the remaining collaborative patterns also have a good number of publications in their credit. This denotes the importance of multidisciplinary teams of scientists working on the virus-related research. These findings are different from those of Hossain,[17] who reported very few single-authored entries on the topic. The review of published literature reveals that more than 10 authors for a single article is not unusual in medical research. This may be due to multiparty collaboration where researchers at the epicenter of the disease outbreak may be seeking outside collaboration. It is important to note here that the WHO made essential recommendations for international cooperation following the Ebola epidemic to avoid failure of therapeutic trials.[19-23] The research institutions at the forefront of an infectious outbreak must be consulted to obtain the benefit of their experience. Of the keywords, COVID-19 was the most frequently used, followed by coronavirus, SARS-CoV-2, pandemic, and pneumonia, as indicated by the findings of Hossain[17] and Hamidah et al.[15]

In order to see the impact of individual articles, a list of the top ten articles was prepared on the basis of their citation count. “Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China,” published by Lancet, achieved the highest number of citations whereas “Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding” was at the bottom of the list. Dehghanbanadaki et al.[16] reported “Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding” as the most cited. Further analysis identified the five articles that obtained more than 1000 citations each, presumably helping accelerate research on COVID-19. The three-factor analysis provides the most productive countries such as China, the USA, and Italy, focusing on the four main keywords including COVID-19, coronavirus, sars-cov-19, and pandemic. The analysis focusing on the map of country collaboration discloses a significant number of publications resulting from collaborative research between the USA and China, which is in agreement with the findings of Hossain.[17] The USA also did much collaborative research with the UK and Canada.

The data summarized in this bibliometric analysis signifies immediate response to an outbreak, both from research and clinical perspective, and its communication. High priority publication of findings would give other countries hints about a possible pandemic and provide them with reaction time to prepare for it. The standard operating procedures of the WHO regarding infectious disease cases also demand communication of such cases as quickly as possible.[22] The same strategy of reporting noteworthy diseases has been followed to successfully combat polio and tuberculosis in most parts of the world.[23]

Another important aspect highlighted by our results is the indigenous nature of the response. Had it been a third world country, we may have seen a delay in reporting, communication, and response to the pandemic. Other possible different outcomes could have been a foreign base of research. For example, two bibliometric reviews published after the Ebola outbreak in West Africa concluded that most of the research, funding agencies, and highly cited articles were based outside Africa.[24,25] While bibliometric analyses are by no means a way of quantifying the actual adequacy of the response on the ground, the apparent effect on limiting the spread and flattening the curve can be correlated with the rapid publication and communication of the data that lead to the formulation of preventive and treatment guidelines.
Limitations and future research directions
This study was limited to WOS publications indexed on the subject area of COVID-19 during 2019–2020. Since our focus was on bibliometric aspects, it was not within the scope of our study to determine whether the incorporated research work was conducted in relevant countries or not. Databases such as PubMed, EMBASE, Google Scholar, Dimension, and Scopus may give different sets of records on searching, but a comparison is out of the scope of this analysis. Future work in this domain would have to verify the present findings with data after the pandemic is over. The studies might also look at the economic and public health impact of individual studies on a thorough quantification of CI and collaboration.

Conclusion
The results of this bibliometric review document are that most of the initial research related to the current coronavirus pandemic was carried out and reported from within the USA and China, with The Journal of Medical Virology and CUREUS being the favorite sources of publications. The research was mostly carried out by large teams. This is an analysis of research done from 2019 to July 2020, so the data and patterns would undoubtedly alter as the virus spreads worldwide. Future studies will provide updates on these dynamics.

Acknowledgment
We would like to thank Institute of Research Promotion (IRP) for English language editing. The manuscript has been read and approved by all the authors, the requirements for authorship have been met, and that all authors believes that the manuscript represents honest work and information is not provided in another form.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References
1. Moorthy V, Henao Restrepo AM, Preziosi MP, Swaminathan S. Data sharing for novel coronavirus (COVID-19). Bull World Health Organ 2020;98:150.
2. Singhal T. A review of coronavirus disease-2019 (COVID-19). Indian J Pediatrics 2020;87:281-6.
3. Shah A, Kashyap R, Tosh P, Sampathkumar P, O’Horo JC. Guide to understanding the 2019 novel coronavirus. Mayo Clin Proc 2020;95:646-52.
4. Wang HJ, Du SH, Yue X, Chen CX. Review and prospect of pathological features of corona virus disease. Fa Yi Xue Za Zhi 2020;36:16-20.
5. Greenland JR, Michelow MD, Wang L, London MJ. COVID-19 infection: Implications for perioperative and critical care physicians. Anesthesiology 2020;132:1346-61.
6. Gross M. Virus outbreak crosses boundaries. Curr Biol 2020;30:R191-4.
7. Bassetti M, Vena A, Giacobbe DR. The novel Chinese coronavirus (2019-nCoV) infections: Challenges for fighting the storm. Eur J Clin Invest 2020;50:e13209.
8. Wong G, Liu W, Liu Y, Zhou B, Bi Y, Gao GF. MERS, SARS, and Ebola: The role of super-spreaders in infectious disease. Cell Host Microbe 2015;18:398-401.
9. Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ, et al. COVID-19: Consider cytokine storm syndromes and immunosuppression. Lancet 2020;395:1033-4.
10. World Health Organization. Coronavirus Disease 2019 (COVID-19): Situation Report, 74; 2020. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports. [Last accessed on 2020 Aug 05].
11. Chiu WT, Huang JS, Ho YS. Bibliometric analysis of severe acute respiratory syndrome-related research in the beginning stage. Scientometrics 2004;61:69-77.
12. Tao Z, Zhou S, Yao R, Wen K, Da W, Meng Y, et al. COVID-19 will stimulate a new coronavirus research breakthrough: A 20-year bibliometric analysis. Ann Transl Med 2020;8:528.
13. Bonilla-Aldana DK, Quintero-Rada K, Montoya-Posada JP, Ramírez-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 Med Infect Dis 2020;33:101566.
14. Chahrouf M, Assi S, Bejani M, Nasrallah AA, Salhab H, Fares M, et al. A bibliometric analysis of COVID-19 research activity: A call for increased output. Cureus 2020;12:e7357.
15. Hamidah I, Sриyono S, Hadha MN. A bibliometric analysis of Covid-19 research using VOSviewer. Indonesian J Sci Technol 2020;5:34-41.
16. Dehghanbanadaki H, Seif F, Yahidi Y, Razi F, Hashemi E, Khoshmirsafta M, et al. Bibliometric analysis of global scientific research on coronavirus (COVID-19). Med J Islamic Republic Iran 2020;34:354-62.
17. Hossain MM. Current Status of global research on novel coronavirus disease (COVID-19): A bibliometric analysis and knowledge mapping [version 1; peer review: awaiting peer review]. F1000Research 2020; 9:374.
18. Lou J, Tian SJ, Niu SM, Kang XQ, Lian HX, Zhang LX, et al. Coronavirus disease 2019: A bibliometric analysis and review. Eur Rev Med Pharmacol Sci 2020;24:3411-21.
19. Keusch G, McAdam KP, Cuff PA, Mancher M, Busta ER. Integrating clinical research into epidemic response: The Ebola experience. Washington, DC: National Academies Press; 2017.
20. World Health Organization. WHO and Civil Society: Linking for Better Health. Geneva: World Health Organization; 2002. Available from: https://www.ncbi.nlm.nih.gov/books/NBK441679/. [Last accessed on 2020 Aug 05].
21. Youde J. Global Health Governance. Malden: Polity; 2012.
22. Parkes MW, Bienen L, Breilh J, Hsu LN, McDonald M, Patz JA, et al. All hands on deck: Transdisciplinary approaches to emerging infectious disease. EcoHealth 2005;2:258-72.
23. Vaz RG, Mkanda P, Banda R, Komkech W, Ekundare-Famiyesin OO, Onyibe R, et al. The role of the polio program infrastructure in response to ebola virus disease outbreak in Nigeria 2014. J Infect Dis 2016;213 Suppl 3:S140-6.
24. Pouris A, Ho YS. A bibliometric analysis of research on Ebola in Science Citation Index Expanded. South Afr J Sci 2016;112:1-6.
25. Yi F, Yang P, Sheng H. Tracing the scientific outputs in the field of Ebola research based on publications in the Web of Science. BMC Res Notes 2016;9:221.