A bibliometric analysis using VOSviewer of publications on COVID-19

Yuetian Yu¹, Yujie Li¹, Zhongheng Zhang², Zhichun Gu³, Han Zhong³, Qiongfang Zha⁴, Luyu Yang⁵, Cheng Zhu⁶, Erzhen Chen⁶

¹Department of Critical Care Medicine, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China; ²Department of Emergency Medicine, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China; ³Department of Pharmacy, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China; ⁴Department of Respiratory and Critical Care Medicine, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China; ⁵Department of Intensive Care Unit, Wuhan Third Hospital, Wuhan University, Wuhan, China; ⁶Department of Emergency Medicine, Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

Contributions: (I) Conception and design: Y Yu, L Yang, C Zhu; (II) Administrative support: Q Zha, Y Li; (III) Provision of study materials or patients: Y Yu, L Yang; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: Y Yu, C Zhu, H Zhong, Z Gu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*These authors contributed equally to this work.

Background: As a global pandemic, COVID-19 has aroused great concern in the last few months and a growing number of related researches have been published. Therefore, a bibliometric analysis of these publications may provide a direction of hot topics and future research trends.

Methods: The global literatures about COVID-19 published between 2019 and 2020 were scanned in the Web of Science collection database. “COVID-19” “Novel Coronavirus” “2019-nCoV” and “SARS-CoV-2” were used as the keywords to reach the relevant publications. VOSviewer was applied to perform the bibliometric analysis of these articles.

Results: Totally 3,626 publications on the topic of COVID-19 were identified and “COVID-19” with a total link strength of 2,649 appeared as the most frequent keyword, which had a strong link to “pneumonia” and “epidemiology”. The mean citation count of the top 100 most cited articles was 96 (range, 26–883). Most of them were descriptive studies and concentrated on the clinical features. The highest-ranking journal was British medical journal with 211 publications and the most cited journal was Lancet with 2,485 citation counts. Eleven articles written by Christian Drosten from Berlin Institute of Virology have been cited for 389 times and 40 articles from Chinese Academy of Sciences have been cited for 1,597 times which are the most cited author and organization. The number of collaborators with China is 44 and the total link strength is 487. The main partners of China are USA, England and Germany. The published literatures have focused on three topics: disease management, clinical features and pathogenesis.

Conclusions: The current growth trends predict a large increase in the number of global publications on COVID-19. China made the most outstanding contribution within this important field. Disease treatment, spike protein and vaccine may be hotspots in the future.

Keywords: Bibliometric analysis; novel coronavirus; coronavirus disease 2019 (COVID-19); SARS-CoV-2; trends

doi: 10.21037/atm-20-4235
View this article at: http://dx.doi.org/10.21037/atm-20-4235
Introduction

As a new acute infectious disease, coronavirus disease 2019 (COVID-19) was first reported in December 2019 in Wuhan, then spread to all the provinces of China and now has become a global pandemic (1). By 20th May 2020, a total of 4,735,622 patients were confirmed in over 200 countries, including 307,537 death cases (2), which resulted in a great public concern. Thus, a series of descriptive researches about the clinical features of COVID-19 have been published by Chinese scholars at the end of 2019. With a deeper understanding of the pathophysiology, more studies focused on antiviral treatment and immune regulation were performed as well as those concentrated on pathology of the disease and vaccine research. In mid-May of this year, more than ten thousand articles have been published and many countries with pandemic have gained a lot of experience from them.

Bibliometrics is a statistical method which could quantitative analysis the research papers concerned about one special topic via mathematical ways (3). It could also access the quality of the studies, analysis the key areas of researches and predict the direction of future studies. The Web of Science (WOS) online database includes almost all the important research papers which also provides built-in analysis tools to produce representative figures. What is more, the search results from WOS could be exported to a software for further analysis like VOSviewer.

However, no bibliometric analysis of publications on COVID-19 has been published till now. As the COVID-19 pandemic has not been fully under control and more knowledge should be obtained from these reference, bibliometric analysis of it is in critical need. Therefore, our study was performed timely to provide a broad understanding of COVID-19 and future research directions.

Methods

The global literatures about COVID-19 published between 2019 to 2020 were scanned in the WOS collection database. The search terms applied to identify the closest matching publication included “COVID-19” or “Novel Coronavirus” or “SARS-CoV-2” or “2019-nCoV” which was used as the keyword in the title. As COVID-19 was first found in Wuhan and a fairly large number of the research papers were written in Chinese, language was not limited during the process of retrieval.

The information for the documents that meet the requirements contained year of publication, language, journal, title, author, affiliation, keywords, document type, abstract and counts of citation which were exported into CSV format. The date of the retrieval was 20th May 2020. VOSviewer (version 1.6.10) was used to analyze the Co-authorship, Co-occurrence, Citation, Bibliographic coupling, Co-citation and themes. Two standard weight attributes are applied which are defined as “Links attribute” and “Total link strength attribute” (4).

Results

Bibliometric analysis of publication output

Totally 15,805 publications on the topic of COVID-19 were identified in WOS database between 2019 and 2020 which included 10,601 (67.1%) original research articles, 1,189 (7.5%) review articles, 2,296 (14.5%) editorials and 1,719 other forms of publications including letters, case reports, etc. Among them, 15,619 (98.8%) papers were published in 2020 (till 20th May) and the other 186 were published in December 2019. Almost all the publications (14,609, 92.4%) were written in English, followed by 623 Chinese publications and 11,575 (73.2%) papers were open access. In all the published papers, 3,626 could be indexed in the WOS core database.

Bibliometric analysis of the keywords

Keywords provided by authors of the paper and occurred for more than 5 times in the WOS core database were enrolled in the final analysis. Of the 4,532 keywords, 344 met the threshold. The keywords that appeared most were “COVID-19” (total link strength 2,649) and “coronavirus” (total link strength 2,024) which had a strong link to “pneumonia” and “epidemiology”. As comparisons of COVID-19, SARS and MERS were another two keywords and the total link strength of each were more than 300 (Figure 1A). A word cloud was also created to show the frequency of the keywords which occurred for more than 10 times. It was indicated that “COVID-19” was the most frequent followed by “pneumonia”, “outbreak” and “infection” (Figure 1B).

Bibliometric analysis of the citations and publications

The top 100 most cited articles in the field of COVID-19 were listed in Table S1. Most of them were clinical studies
Figure 1 Bibliometric analysis of the keywords in publications of COVID-19. (A) Co-occurrence of keywords. The size of nodes indicates the frequency of occurrence. The curves between the nodes represents their co-occurrence in the same publication. The shorter the distance between two nodes, the larger the number of co-occurrence of the two keywords. (B) Word cloud. 137 keywords which occurred for more than 10 times were enrolled. The font size represents the frequency of occurrence. Keywords such as “Coronavirus”, “COVID-19” and “epidemiology” occurred most common. “Prediction” and “Isolation” are rare.
including descriptive studies, case series and case report, the others were research articles which mainly focused on viral genomes and disease transmission. The mean citation count of the top 100 most cited articles was 96 (range, 26–883). All of the papers were published in 2020, and 54 of them were written by Chinese scholars. Thirteen articles were published on the special columns for COVID-19 of the *Lancet*.

Nine hundred and nineteen journals have published papers about COVID-19 and 141 of them have published more than 5 articles. In total, 606 papers were published in the top ten active journal which accounted for 16.7% of the publications in the WOS core database. The highest-ranking journal was *British Medical Journal* (*BMJ*), with 211 publications and an impact factor (IF) of 27.604. The most cited journal was the *Lancet*, with 2,485 citation counts and IF of 59.102 (*Figure 2*).

The top ten most active countries, organizations and authors of COVID-19 publications are listed in Table 1. Eleven articles written by Christian Drosten from Berlin Institute of Virology have been cited for 389 times and 40 articles from Chinese Academy of Sciences have been cited for 1,597 times which are the most cited author and organization. Eight hundred and thirty-eight papers from China have been cited for 7,273 times and the total link strength is 8,162 (*Figure 3*).

**Bibliometric analysis of the co-authorship**

Totally 6,219 authors have participated in the publication of the COVID-19 papers. Among them, Andrei R. Akhmetzhanov from Hokkaido University of Japan has 7 papers which mostly focus on the transmission of COVID-19. The main collaborators with him are Natalie M. Linton and Hiroshi Nishiura from Japan Science and Technology Agency. The total link strength is 49.

Through the domestic and international literature search, it is revealed that 2,037 organizations have published the related papers and 140 of them have over 5 publications.

Huazhong University of Science and Technology has published 90 related papers with 1,268 citations. The main partner of the organization is Wuhan University and most of the researches mainly concentrate on the clinical features of COVID-19. Hongkong University is another important partner with a long-term study on the transmission of diseases.

The number of collaborators with China is 44 and the total link strength is 487 with 838 publications. The main partners of China are USA, England and Germany. Almost as remarkably, the total link strength of Saudi Arabia is 85 and the country cooperates with other 25 countries in the topic of sharing the experience of MERS treatment (*Figure 4*).
**Table 1** The top ten most active countries, organizations and authors of COVID-19 publications

<table>
<thead>
<tr>
<th>Subject</th>
<th>Number of publications</th>
<th>Count of citations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Countries</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>838</td>
<td>7,273</td>
</tr>
<tr>
<td>USA</td>
<td>705</td>
<td>2,102</td>
</tr>
<tr>
<td>England</td>
<td>295</td>
<td>910</td>
</tr>
<tr>
<td>Italy</td>
<td>282</td>
<td>462</td>
</tr>
<tr>
<td>Canada</td>
<td>130</td>
<td>363</td>
</tr>
<tr>
<td>Germany</td>
<td>129</td>
<td>617</td>
</tr>
<tr>
<td>India</td>
<td>128</td>
<td>79</td>
</tr>
<tr>
<td>Australia</td>
<td>114</td>
<td>508</td>
</tr>
<tr>
<td>France</td>
<td>87</td>
<td>242</td>
</tr>
<tr>
<td>Switzerland</td>
<td>86</td>
<td>174</td>
</tr>
<tr>
<td><strong>Organizations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huazhong University of Science and Technology</td>
<td>90</td>
<td>1,268</td>
</tr>
<tr>
<td>Wuhan University</td>
<td>64</td>
<td>1,507</td>
</tr>
<tr>
<td>Hongkong University</td>
<td>56</td>
<td>1,161</td>
</tr>
<tr>
<td>Zhejiang University</td>
<td>47</td>
<td>195</td>
</tr>
<tr>
<td>Fudan University</td>
<td>45</td>
<td>363</td>
</tr>
<tr>
<td>Capital Medical University</td>
<td>43</td>
<td>1,496</td>
</tr>
<tr>
<td>Chinese Academy of Sciences</td>
<td>40</td>
<td>1,597</td>
</tr>
<tr>
<td>Chinese University of Hongkong</td>
<td>40</td>
<td>232</td>
</tr>
<tr>
<td>Harvard Medical School</td>
<td>37</td>
<td>38</td>
</tr>
<tr>
<td>University of Toronto</td>
<td>36</td>
<td>128</td>
</tr>
<tr>
<td><strong>Authors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elisabeth Mahase</td>
<td>32</td>
<td>27</td>
</tr>
<tr>
<td>Gareth Iacobucci</td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td>Abi Rimmer</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>Viroj Wiwanitkit</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Christian Drosten</td>
<td>11</td>
<td>389</td>
</tr>
<tr>
<td>Jiang Shibo</td>
<td>10</td>
<td>52</td>
</tr>
<tr>
<td>Ziad A. Memish</td>
<td>9</td>
<td>115</td>
</tr>
<tr>
<td>Alimuddin Zumla</td>
<td>8</td>
<td>114</td>
</tr>
<tr>
<td>Edward C. Holmes</td>
<td>8</td>
<td>139</td>
</tr>
<tr>
<td>Yang Yang</td>
<td>8</td>
<td>92</td>
</tr>
</tbody>
</table>

**Bibliometric analysis of the bibliographic coupling and co-citation**

The bibliographic coupling map of documents and sources are shown in Figure 5A,B. Seven clusters were obtained from the analysis. Cluster 1 includes 67 items and the research area is clinical features (shown in blue). The representative paper was published in the Lancet in January 2020 by Huang Chaolin. Cluster 7 only has one item which discuss the influence of hypertension and diabetes mellitus on COVID-19 (shown in orange).

Eight clusters of the cited references were obtained by bibliometric analysis. The top three clusters represent the research fields of clinical feature, diseases transmission and treatment which are shown in the colour of red, green and blue. The two biggest clusters of cited sources include 79 and 58 items, respectively. The representative journals are the *Lancet* and *Journal of Virology* (Figure 5C,D).

**Bibliometric analysis of themes and trend topics**

As indicated in Figure 6A, three themes of COVID-19 studies were found. The blue cluster involved clinical trials investigating COVID-19 diagnosis and clinical features. The green cluster involved clinical trials investigating management and emergency preparedness. The red cluster involved clinical trials investigating risk factors and pathogenesis. Figure 6B demonstrates the network map of the trend topics according to the keywords used from December 2019 to April 2020. Indicator shows the current publications from purple to yellow. More studies focused on vaccine, disease treatment and spike protein have been published recently.

**Discussion**

From our current study, 3,626 publications about COVID-19 indexed in WOS core database were analyzed. The published literatures include the following three aspects: disease management, clinical characteristics and pathogenesis. As the most frequent keyword, “COVID-19” has a strong link to “pneumonia” and “epidemiology”. China made the most outstanding contribution within this important field. Disease treatment, spike protein and vaccine may be hotspots in the future.

The COVID-19 epidemic continues to spread around the world till now and the new cases reported outside China...
Figure 3 Bibliometric analysis of the citations. (A) The citations of authors. Eight clusters were shown in different colours. Yuen Kwok-Yung in the purple cluster is the most cited author (191 times); (B) the citations of organizations. Chinese Academy of Sciences in red cluster is the most cited organization (1,036 times); (C) the citations of countries. Different colours indicate different clusters and the size of circles indicates the counts of citations.
Figure 4 Bibliometric analysis of the co-authorship. (A) The co-authorship map of authors which indicates the authors that cooperate in the field of COVID-19 transmission; (B) the co-authorship map of organizations. Huazhong University of Science and Technology has published 51 related papers and cooperates with other 27 institutions; (C) the co-authorship map of countries. The number of collaborators with China is 44 and the total link strength is 290. Different colours indicate different clusters and the size of circles indicates the number of publications. The thickness of the lines represents the link strength of the countries.
Figure 5 Bibliometric analysis of the bibliographic coupling and co-citation. (A) Bibliographic coupling of documents; (B) bibliographic coupling of sources; (C) co-citation of documents; (D) co-citation of sources. Different colors indicate different research areas. The size of the circles represents the counts of co-citations. The distance between the two circles indicates their correlation.
Figure 6 Bibliometric analysis of themes. (A) Distribution of the themes. Three clusters are shown in the map. The red cluster indicates risk factors and pathogenesis. The blue cluster involved clinical trials investigating COVID-19 diagnosis and clinical features. The green cluster represents COVID-19 management and emergency preparedness. (B) Network map of the trend topics according to the keywords used from December 2019 to April 2020. Indicator shows the current publications from purple to yellow. More studies focused on vaccine, drug treatment and spike protein have been published recently. The size of the circles represents the frequency of appearance as the keywords. The distance between the two circles indicates their correlation.
have already exceeded the number of total confirmed cases in Wuhan. The current situation in Europe and America is still very worrisome. To fight against the pandemic, academia joined this “battlefield” as soon as possible to provide recommendations and suggestions to treat the disease. Medical journals with high impact such as the BMJ, Lancet and the New England Journal of Medicine have also opened special columns for COVID-19 (5), which have an advantage in the number of papers published.

COVID-19 is caused by SARS-CoV-2 and like other emerging diseases, the initial focus is often at the clinical characteristics and transmission (6). Thus, as the most commonly used keywords, “COVID-19” and “novel coronavirus” have a strong link to “clinical features” and “epidemiology”. Many articles about the comparison of SARS and MERS were published in order to provide lessons for treatment due to limited knowledge could be obtained at the early stage of the disease (7), and many articles concentrated on clinical features were also most cited (8-10). One of the most important articles was published in 24th January 2020, which demonstrated the clinical characteristics and management of the disease (11). At the same day, person-to-person transmission of COVID-19 in hospital and family settings was verified in another family cluster study which has been cited for more than 300 times till now (12).

As the pandemic was first reported in Wuhan, Chinese scholars wrote almost all the articles published in 2019. With the in-depth study of the disease, more clinical studies were performed in Wuhan and other provinces of China (13). Studies of antiviral therapy like Chloroquine (14), Remdesivir (15), Arbidol (16) and Lopinavir-Ritonavir (17) have been published. However, because of the difference in the severity of the patients and the limited sample size, the results are still controversial (18). Effective drug treatment must be one of the research priorities in the future. As far as vaccines are concerned, monoclonal antibody therapy is still a potential therapeutic intervention to the infectious diseases. Global efforts should be paid on vaccines for COVID-19 and it still has a long way to go (19).

Publications on COVID-19 were retrieved from WOS and the data was analyzed objectively and comprehensively. Nonetheless, some limitations are still inevitable. Firstly, although a large number of new research papers are added to the WOS every day, only a part of them can be indexed in the core database. Thus, most of the non-English language articles were neglected or excluded. As the COVID-19 pandemic originated from Wuhan China, expert consensus written in Chinese might be important and helpful. Secondly, the current growth trends predict a large increase in the number of global publications on COVID-19 which leads to a fairly large number of papers were published in the preprint online database like medRxiv and they were not enrolled in our study.

Conclusions

With the spread of the pandemic, more and more academic papers have been published. It is particularly important to evaluate the quality of such a great number of research papers and obtain valuable information. Scientific and medical research plays a vital role in understanding COVID-19, as well as helping to find solutions to contain its transmission. Effective drug therapy and vaccine research are still future directions.

Acknowledgments

We are grateful to all the medical staffs of our medical rescue team for fighting against COVID-19 together in the last 2 months in Wuhan.

Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-4235). YY serves as an unpaid section editor of Annals of Translational Medicine from Oct 2019 to Sep 2020. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.
References


19. Mukherjee R. Global efforts on vaccines for COVID-19: Since, sooner or later, we all will catch the coronavirus. J Biosci 2020;45:68.
