What We Know about Top 1000 Highly-Cited COVID-19 Papers: A Bibliometric Analysis

Davoud Haseli¹, Heidar Mokhtari², Arezoo Farhadi³, Zahra Shamloo⁴, Mohammad Karim Saberi⁵*, Nafiseh Rezaei⁶

1. Department of Knowledge and Information Science, Faculty of Psychology and Education, Kharazmi University, Tehran, Iran.
2. Department of Library and Information Science, Payame Noor University, Tehran, Iran.
3. Hamadan University of Medical Sciences, Hamadan, Iran.
4. Hearing Disorders Research Center, Loghman Hakim Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
5. Department of Medical Library and Information Sciences, School of Paramedicine, Hamadan University of Medical Sciences, Hamadan, Iran.
6. Ph.D. Student of Medical Library and Information Science. Tehran University of Medical Sciences, Tehran, Iran.

Abstract

Background: Highly cited papers are considered publications with a great impact on a scientific community and have been deeply investigated in different fields.

Aim: This study aimed at analyzing and visualizing the top 1000 highly cited papers on COVID-19.

Methods: As a bibliometric study, this study was conducted by retrieving 1000 highly-cited papers on COVID-19 published during 2019-2021 from Scopus. The search strategy was to obtain 35 related keywords/terms on the COVID-19 as the main term from MeSH and searching them in the fields of paper titles, abstracts, and keywords. Bibliometric techniques such as co-citation analysis, co-authorship analysis and word co-occurrence analysis were used for the study. Data visualization was done by applying the VOSviewer software package and GunnMap.

Results: China, the United States of America (USA), and the United Kingdom (UK) with publishing 418, 353, and 149, mostly cited papers were ranked first-to-third, respectively. The top contributing research institutes were from China and the USA. The top three most productive research institutions were Huazhong University of Science and Technology (N=83), Tongji Medical College (N=76), and Wuhan University (N=64), respectively. The New England Journal of Medicine, the Lancet and JAMA ranked first to third in publishing these papers, respectively. Collaborating countries were mainly of European origin. Research institutes from China, the USA, and the UK had higher collaboration. Keyword clustering showed that the clinical features and laboratory descriptions, risk factors, pathogenic and immunological aspects as well as the managerial aspects and urgent preparation of the disease were topics with high concern and concentration.

Conclusion: This study is the first bibliometric study on the top 1000 highly cited papers on COVID-19 and can be beneficial to researchers in identifying important topics, active producing agents and existing gaps in the literature on the disease. It can be conceived as a reference for COVID-19 researchers and a guide for conducting other bibliometric studies on COVID-19 scientific investigation.

Keywords:
Coronavirus; COVID-19; Bibliometric analysis; Highly-cited papers.

Conflicts of Interest: The Authors declare no conflicts of interest.


This work is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).
Introduction
The outbreak of Coronavirus Disease 2019 (COVID-19) has been declared to be a public health emergency with international concern and recognized as a pandemic disease1. Recently, it is considered as one of the top research topics in the medical field with its increasing trend in research publication, including that of bibliometric studies. Bibliometrics as a tool for measuring scientific impact quantifying research performance2 has been widely used for evaluating scientific research in different fields and a variety of aggregate levels (such as papers, journals, authors, research institutions, research topics, countries/regions, etc.). Bibliometric indicators have been widely and increasingly applied in medical fields3.

Highly cited papers are considered publications with a great impact on a scientific community reflecting different aspects of a scientific discipline. Highly cited papers (especially the top 100 most-cited papers) have been widely investigated different bibliometric aspects in medical and non-medical fields. In medical fields, some related studies can be mentioned, including among others, surgery (4), anesthesia (5), neurosurgery (6), endodontics (7), pediatric neurosurgery (8), radiology (9), traumatic brain injury (10), tuberculosis (11), coronary heart disease 12, gastroenterology and herpetology (13), acute kidney injury (14), neuroimaging (15), endocrinology and metabolism (16), neuroscience (17), pulmonary imaging (18), cardiology (19), microbiology (20), obstetrics and gynecology (21), vaccine and vaccination (22), Hepatitis E virus (23), digestive endoscopy (24), and raninosynostosis (25).

After the outbreak of COVID-19, some bibliometric studies have been conducted on it in the years 2019, 2020, and 2021, investigating the disease from different perspectives (e.g. 23, 26-39). Top highly-cited and most notable papers on COVID-19 have been included in these studies (e.g. 100 highly-cited / most influential articles about COVID-19 (40-42), most notable and highly-disseminated 100 articles on COVID-19 in social media (43, 44), and 50 early-COVID-19 top-cited papers45. However top 1000 highly cited papers on COVID-19 have not been deeply investigated and visualized from a comprehensive bibliometric perspective. This study aimed at analyzing and visualizing these papers with a bibliometric perspective.

Methods
Search strategy and keywords
As a bibliometric study, this study was conducted by retrieving 1000 highly cited papers on COVID-19 published during 2019-2021 from Scopus. The search strategy was to extract 35 related keywords/terms on the topic COVID-19 as the main term from MeSH (Medical Subject Headings) and searching them in the fields of paper titles, abstracts, and keywords. The search query for retrieving related papers was as follows:

Infection" OR "SARS-CoV-2 Infections" OR "COVID-19 Pandemic" OR "COVID 19 Pandemic" OR "COVID-19 Pandemics" OR "Pandemic, COVID-19") AND (LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019))

The search was done on May 2021, and resulted in retrieving 151,276 papers. Out of them, the top 1000 highly-cited papers on COVID-19 limited to the years 2019-2021 were selected and ordered based on their citation numbers, and their bibliometric data were extracted for more analysis and visualization.

**Database used**

We used Scopus for data extraction As the greatest indexing and abstracting database for peer-reviewed scientific literature and commonly-used database for conducting bibliometric studies (46), Scopus is widely scoped than PubMed and Web of Science (WoS), as two other main indexing/abstracting databases (47-49).

**Bibliometric techniques**

Some bibliometric techniques such as co-citation analysis, co-authorship analysis, and word co-occurrence analysis were used for the study. The types, languages, countries of origin, affiliated institutions, and top publishing journals were determined. Co-authorship maps of collaborating countries and research institutes were depicted by co-authorship techniques. Top co-citing journals were identified by co-citation analysis. Word co-occurrence analysis was applied for determining highly frequent keywords and terms and consequent subject clusters.

**Statistical analysis**

Data visualization was done by applying the VOSviewer software package and GunnMap. The former is used for visualizing the citation networks (authors, papers, research institutes, journals, and countries/regions), co-authorship networks (authors, research institutes and countries/regions), co-citation networks (authors, papers, and journals), co-word networks of papers as well as bibliographic couplings and subject clustering (50,51). GunnMap, as a free tool (http://gunnmap.herokuapp.com), was used for creating an infographic data map of the distribution density of papers produced by countries worldwide.

**Results**

**General Information**

The first-ranked highly cited paper with 14172 received citations was authored by 29 authors from China and entitled "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China". The paper was published in The Lancet in 2020. The 1000th highly cited paper was a note with 179 received citations. The mean rate of citations was 528.35 citations per paper (CPP). Out of these top highly cited papers, 13 papers (CPP=355.62), 980 papers (CPP=531.65) and 7 papers (CPP=398.86) belonged to 2019, 2020, and 2021, respectively. These papers included 596 original researches (59.6%, CPP=559.9), 180 reviews (18.0%, CPP=465.6), 117 letters (11.7%, CPP= 533.1), 72 notes (7.2%, CPP=517.4), 25 editorials (2.5%, CPP= 354.2) and 10 other types (1%, CPP=235.8), including short surveys, conference papers and book chapters. 997 papers (CPP=528.1) were in English and only three (CPP=595.3) were in Chinese. 965 papers (96.5%, CPP=539.8) were openly accessed and only 35 papers (3.5%, CPP= 212.9) were not openly-accessed.

**Top most-productive countries, research institutes and journals**

Table 1 shows the top 15 most productive countries, journals, and institutes relating to the top 1000 highly-cited COVID-19 papers. Among contributing countries, the first to the third ranks belonged to China, the USA, and the UK with publishing 418, 353, and 149 papers, respectively. Considering the CPP, Japan (1237.1), Hong-Kong (768.9) and China (684.0) ranked first to third.
Out of the 15 most productive research institutions, 7 and 5 were from China and the USA, respectively. The top three most productive research institutions were Huazhong University of Science and Technology (N=83), Tongji Medical College (N=76), and Wuhan University (N=64). Considering CPPs, two Chinese institutions, i.e. Capital Medical University (1421.1) and Chinese Academy of Sciences (1175.1) were at the top, respectively.

Considering publishing journals, the New England Journal of Medicine with publishing 66 papers (CPP=926.4), the Lancet with publishing 50 papers (CPP= 1293.4), and JAMA (the Journal of the American Medical Association) with publishing 49 papers (CPP=806.0) were first-to-third-ranked journals, respectively. The Science ranked fourth and the Nature and the Journal of Medical Virology ranked fifth in common.

Table 1. Top 15 publishing countries, journals, and research institutes active in contributing to the top 1000 highly-cited papers on COVID-19

<table>
<thead>
<tr>
<th>Category</th>
<th>No. of Pub.</th>
<th>No. of Cit.</th>
<th>CPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>418</td>
<td>285900</td>
<td>684.0</td>
</tr>
<tr>
<td>United States</td>
<td>353</td>
<td>156537</td>
<td>443.4</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>149</td>
<td>73367</td>
<td>492.4</td>
</tr>
<tr>
<td>Italy</td>
<td>104</td>
<td>39070</td>
<td>375.7</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>51</td>
<td>39213</td>
<td>768.9</td>
</tr>
<tr>
<td>France</td>
<td>70</td>
<td>29081</td>
<td>415.4</td>
</tr>
<tr>
<td>Germany</td>
<td>66</td>
<td>38278</td>
<td>580.0</td>
</tr>
<tr>
<td>Canada</td>
<td>57</td>
<td>21496</td>
<td>377.1</td>
</tr>
<tr>
<td>Australia</td>
<td>44</td>
<td>23112</td>
<td>525.3</td>
</tr>
<tr>
<td>Netherlands</td>
<td>39</td>
<td>20795</td>
<td>533.2</td>
</tr>
<tr>
<td>Singapore</td>
<td>39</td>
<td>16254</td>
<td>416.8</td>
</tr>
<tr>
<td>Switzerland</td>
<td>37</td>
<td>16025</td>
<td>433.1</td>
</tr>
<tr>
<td>Spain</td>
<td>34</td>
<td>14059</td>
<td>413.5</td>
</tr>
<tr>
<td>Japan</td>
<td>32</td>
<td>39588</td>
<td>1237.1</td>
</tr>
<tr>
<td>India</td>
<td>22</td>
<td>7138</td>
<td>324.5</td>
</tr>
<tr>
<td>Institution (Country)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huazhong University of Science and Technology (China)</td>
<td>83</td>
<td>70978</td>
<td>855.2</td>
</tr>
<tr>
<td>Tongji Medical College (China)</td>
<td>76</td>
<td>67999</td>
<td>894.7</td>
</tr>
<tr>
<td>Wuhan University (China)</td>
<td>64</td>
<td>63020</td>
<td>984.7</td>
</tr>
<tr>
<td>University of Hong Kong (Hong Kong)</td>
<td>50</td>
<td>34219</td>
<td>684.4</td>
</tr>
<tr>
<td>University of California (United States)</td>
<td>49</td>
<td>21817</td>
<td>445.2</td>
</tr>
<tr>
<td>Harvard Medical School (United States)</td>
<td>46</td>
<td>17386</td>
<td>378.0</td>
</tr>
<tr>
<td>Inserm (France)</td>
<td>37</td>
<td>15035</td>
<td>406.4</td>
</tr>
<tr>
<td>Chinese Academy of Sciences (China)</td>
<td>36</td>
<td>42303</td>
<td>1175.1</td>
</tr>
<tr>
<td>Fudan University (China)</td>
<td>34</td>
<td>16672</td>
<td>490.4</td>
</tr>
<tr>
<td>University of Oxford (United Kingdom)</td>
<td>34</td>
<td>18735</td>
<td>551.0</td>
</tr>
<tr>
<td>Capital Medical University (China)</td>
<td>32</td>
<td>45476</td>
<td>1421.1</td>
</tr>
<tr>
<td>Columbia University (United States)</td>
<td>30</td>
<td>13359</td>
<td>445.3</td>
</tr>
<tr>
<td>Columbia University (United States)</td>
<td>30</td>
<td>13359</td>
<td>445.3</td>
</tr>
<tr>
<td>University of Washington (United States)</td>
<td>29</td>
<td>17014</td>
<td>586.7</td>
</tr>
<tr>
<td>Ministry of Education (China)</td>
<td>28</td>
<td>17182</td>
<td>613.6</td>
</tr>
<tr>
<td>Journal (2019 Impact Factor)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New England Journal of Medicine (74.699)</td>
<td>66</td>
<td>61145</td>
<td>926.4</td>
</tr>
<tr>
<td>The Lancet (60.392)</td>
<td>50</td>
<td>64672</td>
<td>1293.4</td>
</tr>
<tr>
<td>JAMA - Journal of the American Medical Association (45.540)</td>
<td>49</td>
<td>39495</td>
<td>806.0</td>
</tr>
<tr>
<td>Science (41.845)</td>
<td>32</td>
<td>14559</td>
<td>455.0</td>
</tr>
<tr>
<td>Nature (42.778)</td>
<td>24</td>
<td>18784</td>
<td>782.7</td>
</tr>
</tbody>
</table>

This work is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).
Top collaborating countries

Seventy eight countries contributed to publishing the top 1000 highly-cited COVID-19 papers. 56 countries produced at least two and 24 produced at least 10 papers. Figure 1 and Figure 2 depicts the collaboration network of 36 collaborating countries with publishing at least 5 papers in 4 clusters. The sizes of nodes show the frequencies of collaborated published papers and lines among nodes show the collaboration density. The less the distance between two nodes is, the more the collaboration between the two certain countries is. The numbers in parentheses show the number of links a certain country had in collaboration with other countries.

The first cluster (in red) shows the collaboration between 12 European countries (including Italy (31), France (31), Germany (31), Spain (30), Netherlands (25), Denmark (25), Russian Federation (21), Sweden (20), Belgium (19), Greece (18), Norway (17) and Austria (15), two Oceania countries (Australia (29) and New Zealand (17)) and an Asian country (Iran (12)). The second cluster (in green) shows the collaboration of 8 countries (including Canada (31), Singapore (24), Switzerland (22), Brazil (20), Taiwan (17), Turkish (13), South Africa (10) and Vietnam (6)). The third cluster (in blue) includes the collaboration between the USA (35) and five Asian countries (South Korea (24), India (23), Japan (22), Saudi Arabia (17) and Indonesia (7)) and a Central American country (Mexico (14)). The forth cluster (in yellow) highlights the collaboration between China (35) and UK (33), Hong Kong (24), Israel (15), Ireland (13) and Macao (12).

Top collaborating research institutes

1893 research institutes were active in publishing these papers. 544 and 145 institutes published at least 2 and 5 papers. Figure 3 shows the collaboration network of the research institutes of 51 collaborating institutes which published at least 10 papers. Numbers in parenthesis show the number of collaborations of a certain institute with other institutes.

The network consisted of three clusters. The first cluster (in red) included 21 institutes in which 16 USA institutes collaborated with those in Hong Kong, Germany, Italy, Netherlands, and France. The top two collaborating institutes were Harvard Medical School (25) and the Chinese University of Hong Kong (19). In the second cluster (in green), 19 Chinese institutions collaborated. Of them, Huazhong University of Science and Technology and Wuhan University (28 in each) were the top ones. The third cluster (in blue) included 6 British institutes collaborating with 2 institutes from the USA, 2 institutes from Singapore, and 1 institute from Hong Kong.

The first and second ranks in the cluster belonged to the University of Hong Kong (28) and Oxford University (23), respectively.
**Figure 1.** The worldwide distribution of the density of top 1000 highly-cited papers on Covid-19.

**Figure 2.** Collaboration network of countries contributing to top 1000 highly-cited papers on COVID-19

This work is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).
This work is distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 License (CC BY-NC 4.0).
density. The shorter the distance between two nodes are, the higher the co-occurrences between the two related keywords are. The first cluster (in red) included 63 keywords related to clinical laboratory test for diagnosing COVID-19 and its clinical features and descriptions. Including 46 keywords, the second cluster (in green) dealt with risk factors and pathogenic and immunological aspects. The third cluster (in blue) included 27 keywords on COVID-19 managerial aspects and urgent preparation. Most keywords within the second cluster were occurred prior to those within the first and third clusters during the first phases of the outbreak.

**Figure 5.** Key-word co-occurrence network of top 1000 highly-cited papers on COVID-19.

**Discussion**

Our bibliometric study on the top 1000 highly cited papers on COVID-19 shows the current status and trends in research on the disease. The scientific community's response to the disease was immediate as the number of citations received by the papers in these about 2 years after the outbreak clearly shows. The majority of these papers are open access that allows the easy dissemination of needed information among interested researchers and consequent increase in publications and citations. They were mainly original articles focusing on novel approaches and findings that can potentially further our knowledge of the disease. China as the country of origin of the disease and the USA and UK as the two industrialized countries with high contamination has published the majority of these highly-cited papers. The research institutes from China and USA have a main role in publishing these papers. It is needed that the contribution of other countries and research institutes within them, including among others, African countries is encouraged in deep identification and better management of the disease throughout the World.
Well-known and highly prestigious journals with high impact factors (the New England Journal of Medicine, the Lancet, JAMA, Science, and Nature) published the majority of these top-cited papers. These papers were co-cited by highly ranked medical journals, too. However, it is interesting that the Journal of Medical Virology has been very active in publishing and co-citing the papers in line with these journals. This active contribution will increase the impact factor of this journal in future years. Approximately, a quartile (n=245) of top-cited papers were published in the above-mentioned journals.

The majority of highly cited papers on COVID-19 have focused on clinical presentations of the virus and a clear description of the disease as we know little about COVID-19. Other related studies found such a result (40, 41, 45). Potential treatment approaches needed to be investigated in future studies by emphasizing various treatments and efficacy of vaccines.

This study offers important quantitative information on countries, institutions and journals working on the disease. Identifying the most productive countries, institutions, and journals can help potential researchers collaborate with researchers from pioneering countries and institutes and contribute to top journals for making influential works on COVID-19. Published and cited in different journals, the highly cited papers on COVID-19 reflect the complexity of the disease as well as the multidisciplinary nature of research on the disease.

Conclusion
To our knowledge, this study is the first bibliometric study on the top 1000 highly cited papers on COVID-19. We hope that the study is beneficial to researchers in identifying important topics, active producing agents, and existing gaps in the literature on the disease. Despite some limitations, including database selection and citation-based biases, this study can be a reference for COVID-19 researchers and a guide for conducting other bibliometric studies on COVID-19 papers. Interestingly, highly cited papers on COVID-19 are studied from altimetric perspectives for considering their attention in social media for detecting public concerns about the disease.

Acknowledgments
The study was funded by Vice-chancellor for Research and Technology, Hamadan University of Medical Sciences (No. 14000110154).

Conflicts of Interest
The authors declare no conflicts of interest.

Ethics
This study has been ethically approved by the Ethics Committee of Hamadan University of Medical Sciences with code number: IR.UMSHA.REC.1400.006.

Authors' ORCIDs
Heidar Mokhtari
https://orcid.org/0000-0002-2396-8634

Mohammad Karim Saberi,
https://orcid.org/0000-0002-2471-0408

References


33. Atlasi R, Chakoli AN, Ramezani A, Tabatabaei-Malazy O, Larjani B. Scientometric analyzing the output of researchers and organizations on COVID-19 for better conducting the scientific efforts: with a
42. Alghamdi HM. The 100 Most Influential Manuscripts on COVID-19: A Bibliometric Analysis. Available at SSRN 3820550.