The First-Three-Month Review of Research on Covid-19: A Scientometrics Analysis

by Budi Harsanto

Submission date: 16-Sep-2022 08:34AM (UTC+0700)
Submission ID: 1900915366
File name: 25._The_first-three-month_review_of_research_on_COVID-19.pdf (1,006.92K)
Word count: 3897
Character count: 21105

The First-Three-Month Review of Research on Covid-19: A Scientometrics Analysis

Budi Harsanto
Management School
University of Liverpool, UK
Universitas Padjadjaran, Indonesia
harsanto@liverpool.ac.uk
harsanto.id@gmail.com

***Abstract*— Responsiveness is one of dimensions in the concept of responsible research and innovation (RRI). This dimension, among others, is manifested through responses to the challenges faced by society. Responsiveness from the scientific community is clearly evident recently in the response to the recent great challenge of the Covid-19 pandemic. The purpose of this paper is to provide an early review to assess publications in the first three months of the pandemic, since December 2019. Scientometrics and descriptive statistics were used to analyse documents indexed in the Scopus academic database. Open access tools from VOSviewer were used to help in visualising the network of countries and research topic density map. On average, 150.33 documents were published every month, with medicine as the main field of research (62.04%). Researchers from China, US, and UK institutions have published the most documents. Analysis of the keywords shows that the main topics of research regarded acute respiratory diseases, contact tracing, and molecular epidemiology. Although fields such as psychology, mathematics, computer science, engineering, nursing, business, management and accounting still have very few Covid-19 related studies at the moment, these fields likely to contribute in the future with regard to the great impact of this pandemic.**

Keywords— review, scientometrics, Covid-19, VOSviewer

I. INTRODUCTION

Responsible research and innovation (RRI) has several dimensions including anticipation, reflexivity, inclusion, and responsiveness [1]. All these dimensions are interconnected. RRI applies both in normal and crisis situations [2]. In crisis situations, various dimensions of RRI provide a framework for consideration in the research and innovation practice. By taking the example of the Zika crisis, Monteiro et. al. highlighted the need for care in producing knowledge in a transparent and participative manner, to be able to make a positive contribution to society [2]. With a scale far greater than Zika, the world community is currently facing challenges in the form of the Covid-19 pandemic. Referring to WHO, Covid-19 is "an infectious disease caused by a newly discovered coronavirus" [3]. The disease first emerged in December 2019 in China and has now spread to almost all countries in the world. Previously its status was an epidemic, and on 11 March 2020, its status was raised become a pandemic [4].

In addressing these challenges, among the most responsive is the scientific community. This is marked by the abundance of scientific studies from various authors from various fields from various affiliates around the world to understand and find solutions to this pandemic. This is also marked by a rare event where dozens of leading publishers are opening their archives to the public for free immediate access to all publications as well as available data regarding Covid-19 [5]. This responsiveness is beneficial for the public in terms of emergency responses to this ongoing pandemic, given that access to these resources is not cheap. One of the representatives of the publishers revealed that, "open access and open science are critical, especially in times of crisis" [5]. Open access, openness, and transparency as well as responsiveness to societal challenges are some of the elements discussed in RRI, especially in the responsiveness dimension [1]. The abundance of studies on Covid-19 are indexed in academic databases to facilitate effective research and analysis for various purposes.

This paper aims to conduct a scientometrics analysis of the publications on Covid-19 in the first three months of the outbreak, using data in certain academic databases with wide coverage. Scientometrics analysis can help provide a reliable examination of the trends, current status, and dynamics of a scientific field [6]. In this way, researchers, funding bodies, policy makers, and the general public can be informed on certain knowledge and indications about the network contained therein. This paper is structured as follows. The first section discusses research background. The second section explain a detailed methodology to ensure transparency, rigorousness and enable replicability. The third section provide results and discussion.

II. METHODOLOGY

In this paper, Scopus academic database from Elsevier is used as a source of literature search. This academic database was chosen because its coverage is broader than the Web of Science from Thomson Reuters. This database was also chosen to find out the research landscape of various fields, not only in biomedical fields such as PubMed. The search was conducted on 25 March 2020 using search terms, including: (corona virus) OR (2019 nCoV) OR (SARS CoV 2) OR (Covid 19) in the article title, abstract and keywords. This rough search resulted in 1,960 articles. Refinement is performed by checking the year of publication. Publications other than 2019 and 2020 were excluded from the search and resulted in 602 articles. Then to obtain more accurate results, the search results narrow it down again using the same search term but in the title. This process resulted in 451 articles.

2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)

Two types of data are then exported from Scopus, namely main information (citation information, bibliographical information, abstract and keywords) and refine values (information from refinement tools in Scopus); both in CSV format. To visualise bibliometric networks, main information from Scopus is processed by VOSviewer software version 1.6.14 (released January 2020) [7]. Analysis conducted using the VOSviewer was the analysis of co-authorship with units of analysis countries. The counting method used is 'fractional counting' where the weights of a link is not same, but fractionalised [8]. Previous researchers revealed that fractional counting is preferable than full counting either based on theoretical and empirical consideration [9]. Minimum number of documents per country was five documents with the minimum number of citations zero. Verification is done manually to ensure data accuracy by including non-countries terms in the analysis. Furthermore, co-occurrence of keywords also analysed using VOSviewer with the minimum number of occurrences more than or equal two; of the 460 keywords, 84 meet the threshold. Other statistical analyses were carried out using Microsoft Excel. For impact factor data, Journal Citation Reports 2018 data from the Clarivate Analytics was used [10].

III. Results and Discussion

Of the 451 publications found, almost all 449 articles (99.56%) were published in journals (Table 1). When averaged, 150.33 articles per month were published in the first three months of this outbreak.

TABLE I. Document Type, Access Type, Year and Language

Document type	Number of papers	Percent
Source type		
Journals	449	99.56
Conference Proceedings	2	0.44
Document type		
Article	199	44.12
Letter	79	17.52
Editorial	68	15.08
Note	68	15.08
Review	26	5.76
Short Survey	5	1.11
Erratum	4	0.89
Conference Paper	2	0.44
Access type		
Open Access	388	86.03
Other	63	13.97
Year		
2020	446	98.89
2019	5	1.11
Language		
English	439	96.48
Chinese	4	0.88
German	3	0.66
Spanish	3	0.66
French	2	0.44
Italian	2	0.44
Dutch	1	0.22
Icelandic	1	0.22

The vast majority of documents published in the journals are article journals, followed by other types of documents such as letters and editorials. Given that so many leading publishers have opened access to this topic, it is not surprising that more than 85% of publications are open access. Nearly 99% of papers were published in 2020.

Exploration of the distribution of publications based on scientific fields was then carried out based on the subject-area categorisation in Scopus (Table 2). The results are not surprising where the majority of publications (62.04%) came from the field of medicine. There are also a large number of publications from the immunology and microbiology categories (16.43%) which indicate an emphasis on this specific area in studying the outbreak. Other subject area such as psychology, mathematics, computer science, engineering, nursing, business, management and accounting are still very few at the moment, but they are likely to contribute to the future with regard to the great impact of the pandemic.

TABLE II. Scientific Field

Subject Area	No of papers	Percent
Medicine	389	62.04
Immunology and Microbiology	103	16.43
Biochemistry, Genetics and Molecular Biology	33	5.26
Pharmacology, Toxicology and Pharmaceutics	20	3.19
Neuroscience	12	1.91
Agricultural and Biological Sciences	8	1.28
Environmental Science	8	1.28
Mathematics	6	0.96
Multidisciplinary	6	0.96
Psychology	5	0.80
Undefined	5	0.80
Computer Science	4	0.64
Dentistry	4	0.64
Health Professions	4	0.64
Engineering	3	0.48
Nursing	3	0.48
Social Sciences	3	0.48
Veterinary	3	0.48
Chemistry	2	0.32
Physics and Astronomy	2	0.32
Business, Management and Accounting	1	0.16
Chemical Engineering	1	0.16
Energy	1	0.16
Materials Science	1	0.16

To obtain a deeper insight, the contribution of countries where author's affiliates are based is analysed based on the number of documents, citations, and total link strength (Table 3). Based on the number of documents, the three top countries are China (158 papers), USA (78) and UK (48). Based on the number of citations, the three top countries are China (262 citations), US (125), and Germany (99). The majority of countries in the top 10 contributing countries are from Europe (5 countries), followed by Asia (3 countries), and North America (2 countries). In total link strength, the top three countries are the USA (96), UK (80), and China (77). Total link strength measures the strength of links in inter-countries collaboration. The higher the value, the stronger the link [11].

TABLE III. TOP CONTRIBUTING COUNTRIES

Rank	Country	No of papers	Citations	Total link strength
1	China	158	262	51.00
2	United States	78	125	46.00
3	United Kingdom	48	99	25.00
4	Canada	23	18	18.00
5	Germany	23	114	17.00
6	Saudi Arabia	18	47	16.00
7	France	25	35	14.00
8	Italy	26	50	13.00
9	The Netherlands	11	28	9.00
10	Singapore	16	10	8.00

The density maps based on total link strength shown in Figure 1. This density map focuses on countries with a major contribution marked by contributions of at least five articles (default in VosViewer). From 102 countries, there are 25 countries that fall into these criteria. There are five clusters formed from links between these countries. Cluster 1 consists of Australian and six European countries (Denmark, France, Germany, Netherlands, Spain, and Sweden). Cluster 2 consists of countries in America (Brazil, Canada), Europe (Italy and Switzerland), Africa (South Africa), and Asia (Taiwan). Cluster 3 consists of China, Japan, Singapore, UK, and US. The three countries with with highest number of papers, citations and total link strength is in cluster 3. Cluster 4 consists of three countries in Asia (India, South Korea, and Thailand). Cluster 5 consists of two countries in Middle East (Egypt and Saudi Arabia).

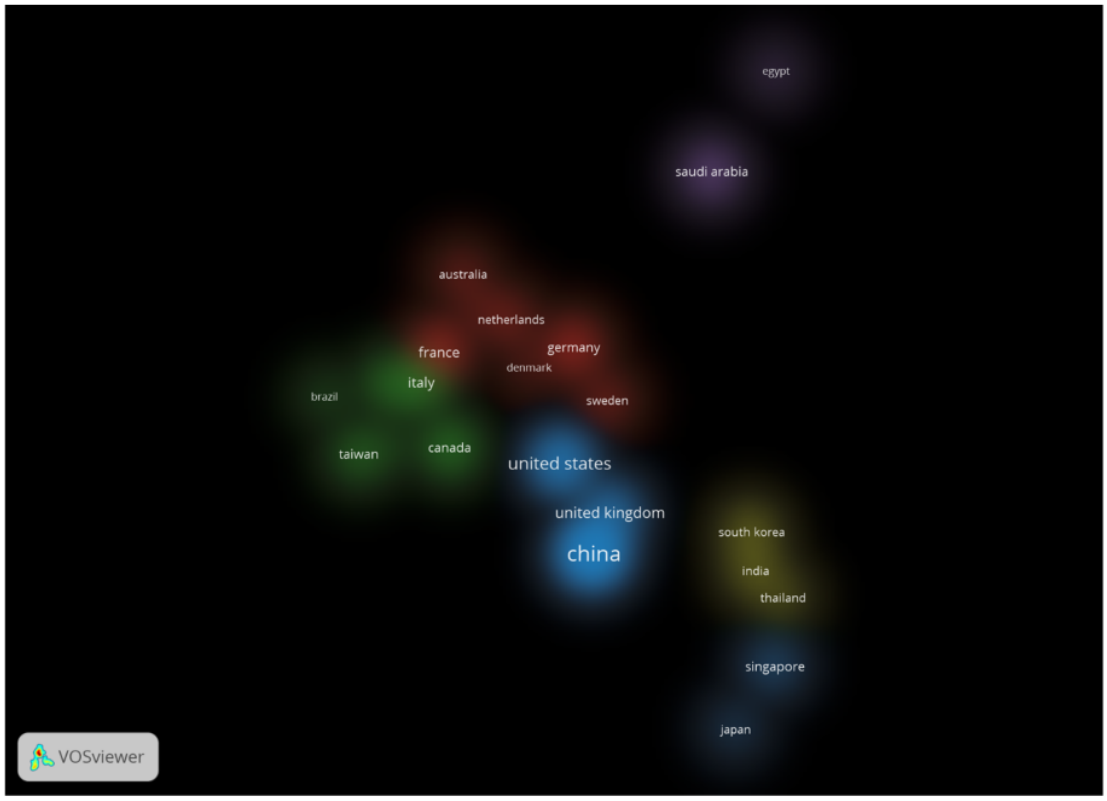

Fig. 1. Density map contributing countries based on fractional counting

Top contributing funding sponsors are from China, North America, and Europe (Table V). Funding sponsors from China contributed to more than 40 papers. The three top funding bodies from China are the National Natural Science Foundation of China (28 papers), followed by the Chinese Academy of Sciences and the National Basic Research Program of China (973 Programs) (each 5 papers). Funding sponsors from North America include the Canadian Institutes of Health Research and Centers for Disease Control and Prevention (each 4 papers), followed by the Bill and Melinda Gates Foundation and the National Institutes of Health (each 3 papers). Funding sponsors from Europe include the Horizon 2020 Framework Program (4 papers) and the European and Developing Countries Clinical Trials Partnership (3 papers).

Data on funding sponsors were identified and aggregated by Scopus. It is important to note that the research funding system is different from country to country. For example, countries in European use funding originating from the European Union, not from national funding bodies like other countries.

TABLE IV. FUNDING SPONSORS

Rank	Funding Sponsors	No of papers	Per cent
1	National Natural Science Foundation of China	28	5.15
2	Chinese Academy of Sciences	5	0.92
3	National Basic Research Program of China (973 Program)	5	0.92
4	Canadian Institutes of Health Research	4	0.74
5	Centers for Disease Control and Prevention	4	0.74
6	Chinese Academy of Medical Sciences	4	0.74
7	Horizon 2020 Framework Programme	4	0.74
8	Bill and Melinda Gates Foundation	3	0.55
9	European and Developing Countries Clinical Trials Partnership	3	0.55
10	National Institutes of Health	3	0.55

The ten top journals that became publication outlets are presented in table VI. Two of these journals had very high impact factors >50, namely Lancet (59,102) and JAMA Journal of The American Medical Association (51,273). Two other journals with a high impact factor (> 15) are BMJ Clinical Research Ed (27,604) and Intensive Care Medicine (18,967). In addition to the highest impact factor, Lancet is also become the top source title on this topic (51 papers). The next top publication outlets are BMJ Clinical Research Ed (28 papers) and Journal of Medical Virology (27 papers).

TABLE V. TOP SOURCE TITLE

Rank	Source Title	Impact Factor	No of papers	Per cent
1	Lancet	59.102	51	11.75
2	BMJ Clinical Research Ed	27.604	28	6.45
3	Journal of Medical Virology	2.049	27	6.22
4	Euro Surveillance Bulletin Europeen Sur Les Maladies Transmissibles European Communicable Disease Bulletin	NA	21	4.84
5	JAMA Journal of The American Medical Association	51.273	15	3.46
6	Intensive Care Medicine	18.967	11	2.53
7	BMJ	2.376	10	2.30
8	Travel Medicine and Infectious Disease	4.868	10	2.30
9	Emerging Microbes and Infections	6.212	8	1.84
10	Canadian Journal of Anesthesia	3.374	6	1.38

The top authors found are Rothe et. al. with a total of 46 citations [10] (Table VII). This number of citations is the citations in the Scopus database. When compared with other databases, the number of citations of the paper on Google Scholar is far greater, reaching 249 citations; while on the Web of Science, there were fewer: 19 citations. Rothe et.al. paper investigates the transmission of viruses in Germany, published 2020 in the New England Journal of Medicine. Hui et. al. and Wu, Leung and Leung became the next top authors with citations of 40 and 37 respectively, both published in 2020 [11], [12]. Three journals that contain more than one top cited papers are the International Journal of Infectious Diseases, The Lancet, and the Journal of Medical Virology (each 2 papers). The aspects of transmission and investigation to understand the characteristics of viruses and diseases were the ones discussed by the most cited papers. Except for Cell Research journals, all journals that are the most cited papers publication outlets are found in the top source title (Table VI).

TABLE VI. THE MOST CITED PAPERS

Rank	Authors, Year and Title	Journal	Citations
1	Rothe et. al (2020), Transmission of 2019-NCOV infection from an asymptomatic contact in Germany [12]	New England Journal of Medicine	46
2	Hui et. al. (2020), The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — The latest 2019 novel coronavirus outbreak in Wuhan, China [13]	International Journal of Infectious Diseases	40
3	Wu, Leung and Leung (2020), Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study [14]	The Lancet	37
4	Ji et. al. (2020), Cross-species transmission of the newly identified coronavirus 2019-nCoV [15]	Journal of Medical Virology	26
5	Wang et. al. (2020), Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro [15]	Cell Research	21
6	Wu & McGoogan (2020), Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases from the Chinese Center for Disease Control and Prevention [16]	JAMA - Journal of the American Medical Association	20
7	Corman (2020), Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR [17]	Euro surveillance: bulletin Europeen sur les maladies transmissibles	19
8	Chen et. al. (2020), Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records [18]	The Lancet	18
9	Zhao et. al. (2020), Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak [19]	International Journal of Infectious Diseases	16
10	Wang, Tang, and Wei (2020), Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China [20]	Journal of Medical Virology	12

Finally, for research topics, paper keywords are analysed using VOSviewer. In addition to identifying the topics that appear most often, the analysis also includes the relationship between one keyword with other keywords. Of 460 keywords, keywords with an occurrence of at least 2 times consist of 84 keywords. The processing results show 13 unique clusters of keywords. The three highest clusters mainly discuss acute respiratory diseases (13 items of keywords), followed by contact tracing (11 items), and molecular epidemiology (10 items). In addition to keywords used as search strings in searches, the most prominent keywords appear are pneumonia (occurrence 18 times with a total link strength of 68), followed by epidemiology (14 times, total link strength of 59).

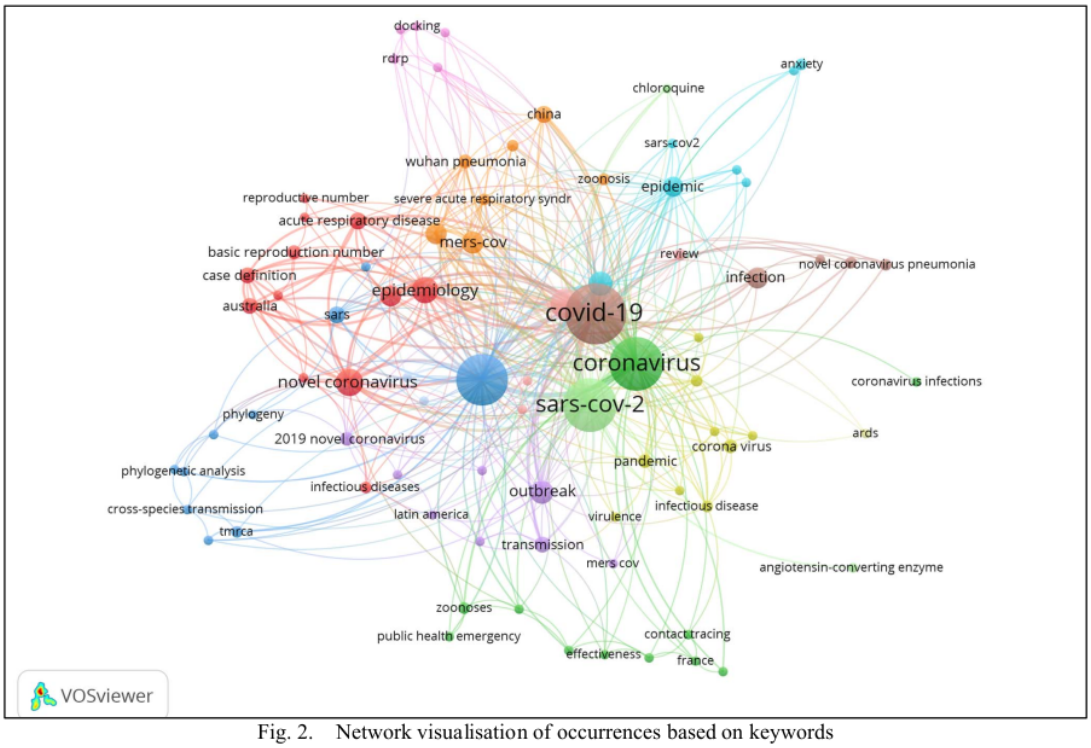

Fig. 2. Network visualisation of occurrences based on keywords

IV. Conclusion

This paper aims to analyse the publication of Covid-19 in the first three months of the outbreak using the scientometrics approach. The analysis shows that most publications were from journals, especially peer-reviewed articles. The majority of publications come from the medicine, immunology and microbiology, and biochemistry, genetics, and biology fields. Based on country contributions, there are five clusters formed. The three countries with the highest contributions are China, US, and UK. Top publication outlets are The Lancet, followed by BMJ Clinical Research Ed and Journal of Medical Virology. Analysis of keywords shows that there are 13 unique clusters with three main clusters discuss acute respiratory diseases, contact tracing and molecular epidemiology.

This paper has several limitations. The timeframe analysed was limited to the first three months of the outbreak. This study could be a good foundation for knowing researchers' responsiveness as well their affiliations, funding bodies, and countries in the early outbreak period. However, to get a complete picture of the development of the scientific landscape regarding the Covid-19 phenomenon, a similar analysis can be carried out after this pandemic end, which hopefully end soon. Analysis that has not been performed in this study can also be performed in the future research, for example, co-authorship analysis based on authors and organisations. Further studies can also discuss deeper the findings based on particular taxonomy or perspective (for example, RRI or disaster management or other fields).

Acknowledgment

The author would like to thank three anonymous reviewers for their valuable feedback on earlier versions of this paper, especially regarding full and fractional counting comparison.

References

[1] J. Stilgoe, R. Owen, and P. Macnaghten, "Developing a framework for responsible innovation," *Res. Policy*, vol. 42, no. 9, pp. 1568–1580, 2013.

[2] M. Monteiro, C. Shelley-Egan, and J. Dratwa, "On irresponsibility in times of crisis: learning from the response to the Zika virus outbreak," *J. Responsible Innov.*, vol. 4, no. 1, pp. 71–77, 2017.

[3] WHO, "Coronavirus." [Online]. Available: https://www.who.int/health-topics/coronavirus#tab=tab_1. [Accessed: 26-Mar-2020].

[4] WHO, "WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020." [Online]. Available: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020. [Accessed: 26-Mar-2020].

[5] Wellcome, "Publishers make coronavirus (COVID-19) content freely available and reusable | Wellcome." [Online]. Available: https://wellcome.ac.uk/press-release/publishers-make-coronavirus-covid-19-content-freely-available-and-reusable. [Accessed: 26-Mar-2020].

[6] P. Vezyridis and S. Timmons, "Evolution of primary care databases in UK: A scientometric analysis of research output," *BMJ Open*, vol. 6, no. 10, 2016.

[7] Centre for Science and Technology Studies, Leiden University, "VOSviewer." 2020.

[8] N. J. van Eck and L. Waltman, *VOSviewer manual*, no. January. 2013.

[9] A. Perianes-Rodriguez, L. Waltman, and N. J. van Eck, "Constructing bibliometric networks: A comparison between full and fractional counting," *J. Informetr.*, vol. 10, no. 4, pp. 1178–1195, 2016.

[10] Clarivate Analytics, "Journal Citation Reports," 2018. [Online]. Available: https://jcr.clarivate.com. [Accessed: 26-Mar-2020].

[11] N. J. Van Eck and L. Waltman, "VOSviewer Manual," *1 January 2013*, no. January, pp. 1–28, 2013.

[12] C. Rothe *et al.*, "Transmission of 2019-nCoV infection from an asymptomatic contact in Germany," *N. Engl. J. Med.*, 2020.

[13] D. S. Hui *et al.*, "The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—The latest 2019 novel coronavirus outbreak in Wuhan, China," *Int. J. Infect. Dis.*, vol. 91, pp. 264–266, 2020.

[14] J. T. Wu, K. Leung, and G. M. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study," *Lancet*, vol. 395, no. 10225, pp. 689–697, 2020.

[15] W. Ji, W. Wang, X. Zhao, J. Zai, and X. Li, "Cross species transmission of the newly identified coronavirus 2019 nCoV," *J. Med. Virol.*, vol. 92, no. 4, pp. 433–440, 2020.

[16] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention," *Jama*, 2020.

[17] V. M. Corman *et al.*, "Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR," *Eurosurveillance*, vol. 25, no. 3, 2020.

[18] H. Chen *et al.*, "Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records," *Lancet*, vol. 395, no. 10226, pp. 809–815, 2020.

[19] S. Zhao *et al.*, "Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak," *Int. J. Infect. Dis.*, vol. 92, pp. 214–217, 2020.

[20] W. Wang, J. Tang, and F. Wei, "Updated understanding of the outbreak of 2019 novel coronavirus (2019 nCoV) in Wuhan, China," *J. Med. Virol.*, vol. 92, no. 4, pp. 441–447, 2020.

The First-Three-Month Review of Research on Covid-19: A Scientometrics Analysis

ORIGINALITY REPORT

9%	12%	10%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	researchrepository.ucd.ie Internet Source	3%
2	Www.Medrxiv.Org Internet Source	2%
3	e-space.mmu.ac.uk Internet Source	2%
4	flexikon.doccheck.com Internet Source	2%

Exclude quotes	Off	Exclude matches	< 2%
Exclude bibliography	Off		