 doi:10.20944/preprints202006.0299.v1

Article

A Bibliometric Analysis of COVID-19 across Science and Social Science Research Landscape

Aleksander Aristovnik *, Dejan Ravšelj and Lan Umek

Faculty of Public Administration, University of Ljubljana, 1000 Ljubljana, Slovenia; dejan.ravselj@fu.uni-lj.si; lan.umek@fu.uni-lj.si

* Correspondence: aleksander.aristovnik@fu.uni-lj.si

Abstract: The COVID-19 pandemic caused by the novel coronavirus emerged in Wuhan City, Hubei province of China at the end of 2019, has radically transformed the lives of people around the world. Due to its fast spreading, it is currently considered as a worldwide health, social and economic concern. The lack of knowledge on this area has encouraged academic sphere for extensive research, which is reflected in exponentially growing scientific literature in this area. However, current state of COVID-19 research reveals only early development of knowledge, while a comprehensive and in-depth overview remains neglected. Accordingly, the main aim of this paper is to fill the aforementioned gap in the literature and provide an extensive bibliometric analysis of COVID-19 research across science and social science research landscape. The bibliometric analysis is based on the Scopus database including all relevant and latest information on COVID-19 related publications (n=10,344) in the January-May 2020 period. The findings emphasize an importance of a comprehensive and in-depth approach considering different scientific disciplines in COVID-19 research. The understanding of the evolution of emerging scientific knowledge on COVID-19 is beneficial not only for scientific community but also for evidence-based policymaking in order to prevent and address the COVID-19 pandemic.

Keywords: COVID-19; coronavirus; pandemic; science; social science; bibliometric analysis

1. Introduction

Since 2000s, the world has witnessed two large-scale disease outbreaks. These are Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS), which emerged in 2003 and 2012, respectively, and caused a worldwide pandemic that claimed thousands of human lives [1]. In December 2019, a new strain of coronavirus (COVID-19), not previously identified in humans, has emerged in Wuhan City, Hubei province of China. The virus has begun to spread exponentially across all inhabited continents and the number of cases and deaths related to COVID-19 has soon exceeded the numbers of other two coronaviruses (SARS and MERS). The outbreak of COVID-19 is a typical public health emergency. Its high infection rate makes it a huge threat to global public health [2-4]. However, its rapid spread has not only affected the lives of many people around the world, but also disrupted the pattern of social and economic development, leading to incalculable social and economic losses [5]. In almost 6 months, more than 9 million cases and more than 470,000 deaths have been seen at the global level [6]. International institutions have therefore announced the global economy is now in a recession – as bad or worse than in the global financial crisis of 2009, arguing this recession will affect both developed and developing countries [7,8]. Therefore, it is not surprising, why the COVID-19 pandemic has attracted the attention of the academic sphere and spurred a new wave of research in this area.

The recent bibliometric studies considering broader aspect of coronavirus research in time provide some interesting findings. Taking into account previous coronavirus pandemics Hu et al. [9] establish that the highest research interest occur in the first year after outburst. This is further

 doi:10.20944/preprints202006.0299.v1

confirmed by the study addressing coronavirus research trends during the last 50-years period [10]. Therefore, it is not surprising why recently COVID-19 has become the central topic in the recent scientific literature, since the research addressing various aspects of COVID-19 may be the key to mitigating the current COVID-19 pandemic as well as their consequences [11]. In the literature, there exist also several recent bibliographic studies, which are focusing only on COVID-19 research, revealing that China and the United States have the largest COVID-19 scientific production [12-16]. Moreover, it is established that virology, epidemiology, clinical features, laboratory examination, radiography, diagnosis and treatment are the current research hotspots of COVID-19 [14,15]. Finally, the majority of published documents on COVID-19 are published in prestigious journals with high impact factors, including the Lancet, BMJ – Clinical Research Ed. and Journal of Medical Virology [12,16].

Although, the absence of knowledge on the novel COVID-19 has grabbed the attention of the academic sphere, spurring a new wave of research into the virus, yet, the vast majority of recent studies chiefly consider health-related issues, leaving other aspects neglected, as indicated by the latest literature. Moreover, COVID-19 research's current status is only of the early development of knowledge. Therefore, the literature stresses that greater research should be conducted in less-explored areas, including life, physical and social sciences [14]. Accordingly, the main aim of this paper is to provide an extensive bibliometric analysis on COVID-19 research in first five months of 2020. Although there already exist several papers addressing bibliometric analysis of COVID-19 research, several research gaps are identified, which are carefully tackled by this paper. First, the existing bibliometric studies are predominantly focused on general analysis of COVID-19 research, showing the importance of health sciences in this area, while detailed insight considering different research landscapes remain neglected. Therefore, this paper provides in-depth bibliometric analysis by considering various science and social science research landscapes. Second, the predominant part of the research are mostly addressing databases containing document information only. Accordingly, this paper extends the analysis on a comprehensive database including document and source information, allowing the bibliometric analysis in different research landscapes. Finally, recent research is neglecting also some sophisticated bibliometric approaches. Therefore, this paper provides an innovative approach, allowing showing all possible logical relations between different research landscapes.

Thus, the main aim of this paper is to provide an unprecedented, comprehensive and in-depth examination of COVID-19 research across different research landscapes, which can suggest important guidelines for researchers about the avenues for future research. The remaining sections of this paper are structured as follows. The second section presents materials and methods. In the third section, the results are discussed. The paper ends with conclusion, where main findings are summarized.

2. Materials and Methods

A comprehensive bibliometric data on COVID-19 related research is obtained throughout two consecutive phases. The first phase involves identification of all relevant documents or publications on June 1, 2020 in the Scopus database on document information. The applied search query includes a wide range of COVID-19 related keywords: "novel coronavirus 2019", "coronavirus 2019", "COVID 2019", "COVID19", "COVID 19", "COVID-19", "SARS-CoV-2", "HCoV-19", "2019-nCoV" and "severe acute respiratory syndrome coronavirus 2". Additionally, the search query is restricted to the current year 2020 and English language. The second phase involves supplementing the presented Scopus database on document information with Scopus CiteScore metrics containing source related information (e.g. citations, rankings, SNIP, etc.). The process of obtaining and merging the relevant data is facilitated by Python programming language.

Then, an in-depth bibliometric analysis followed, allowing for an innovative approach to literature review. Namely, the structured literature review represents a traditional approach to analyse and review scientific literature, providing an in-depth overview of the content. However, this approach suffers from several limitations related to subjective factors, time-consumption and

 doi:10.20944/preprints202006.0299.v1

efficiency. The application of modern bibliometric approaches reduces the aforementioned limitations and provide an effective way to handle extensive collections of scientific literature [17]. Therefore, an innovative quantitative bibliometric analysis is conducted to assess the state of current COVID-19 research across different research landscapes by innovative statistical approaches using Python programming language and software bibliometric tool VOSviewer.

3. Results

An overview of scientific documents utilised in this study is presented in Table 1. A total of 10,344 documents written by 44,439 different authors and published in 1,978 journals were utilised in this study, whereby 3,790 (37%) of them have at least one citation in the Scopus database providing a total of 48,044 citations. For these documents, the average citations per document were 12.68 and the average authors per document were 4.30. A major proportion of the documents were articles (39%) and letters (27%). Much smaller proportion of the documents were editorials (11%) notes (10%) and reviews (10%). Finally, there was a negligible proportion of other documents (3%) such as short surveys, conference papers, erratums and data papers.

Table 1. Overview of scientific documents on COVID-19 research (January-May 2020).

Database summary	Findings
Bibliometric items	**Number**
Total documents	10,344
Total authors	44,439
Total journals	1,978
Total citations	48,044
Cited documents	3,790
Average citations	12.68
Average authors	4.30
Document type	**Number (share)**
Article	4,001 (39%)
Letter	2,827 (27%)
Editorial	1,158 (11%)
Note	1,024 (10%)
Review	1,017 (10%)
Other	317 (3%)

Source: Authors' elaboration based on Scopus database, June 2020.

Table 2 presents most relevant (top 20) journals in COVID-19 research by number of documents. They contain about one-fifth (19%) of total documents and cover more than half (53%) of total citations. Most of them are ranked into the first quartile (Q1) and have a relatively high source normalized impact per paper (SNIP). Moreover, the majority of these journals are subject to health sciences and they are classified predominantly in the following subsubject areas: infectious diseases, general medicine and microbiology (medical). These findings are in line with previous COVID-19 related bibliometric research. However, all of the existing bibliometric studies are neglecting the fact that sciences are strongly intertwined, leading to a lack of understanding of the COVID-19 research in other research landscapes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 June 2020 doi:10.20944/preprints202006.0299.v1

Table 2. Most relevant journals in COVID-19 research (January-May 2020).

Source title	Number of documents	Number of citations	Sub-subject area (ranking) 2019	SNIP 2019
Journal of Medical Virology	221	1,781	Infectious Diseases (108/283, Q2) Virology (37/66, Q3)	0.780
The Lancet	196	7,493	General Medicine (1/529, Q1)	21.313
BMJ Case Reports	187	573	General Medicine (289/529, Q3)	0.364
JAMA - Journal of the American Medical Association	121	4,077	General Medicine (4/529, Q1)	11.131
Journal of Infection	121	429	Infectious Diseases (21/283, Q1) Microbiology (medical) (13/115, Q1)	1.587
Infection Control and Hospital Epidemiology	108	68	Infectious Diseases (91/283, Q2) Microbiology (medical) (39/115, Q2) Epidemiology (40/93, Q2)	1.358
Travel Medicine and Infectious Disease	97	300	Public Health, Environmental and Occupational Health (73/516, Q1) Infectious Diseases(82/283, Q2)	1.184
International Journal of Infectious Diseases	97	766	Infectious Diseases (59/283, Q1) Microbiology (medical) (26/115, Q1)	1.426
The Lancet Infectious Diseases	91	1,103	Infectious Diseases (4/283, Q1)	7.234
International Journal of Environmental Research and Public Health	90	169	Public Health, Environmental and Occupational Health (174/516, Q2) Health, Toxicology and Mutagenesis (68/128, Q3) Pollution (58/120; Q2)	1.248
Science of the Total Environment	88	164	Environmental Engineering (10/132, Q1) Pollution (13/120, Q1) Waste Management and Disposal (10/100, Q1) Environmental Chemistry (17/115, Q1)	1.977
Head and Neck	78	73	Otorhinolaryngology (5/106, Q1)	1.356
Medical Hypotheses	75	16	General Medicine (99/529, Q1)	0.509
Journal of Clinical Virology	70	44	Infectious Diseases (44/283, Q2) Virology (19/66, Q2)	1.238
Psychiatry Research	66	109	Psychiatry and Mental health (154/506, Q2) Biological Psychiatry (25/38, Q3)	0.968
Brain, Behavior, and Immunity	64	248	Behavioral Neuroscience (2/73, Q1) Immunology (27/200, Q1) Endocrine and Autonomic Systems (3/24, Q1)	1.416
New England Journal of Medicine	62	6,049	General Medicine (2/529, Q1)	13.212
Science	62	625	Multidisciplinary (2/111, Q1)	7.521
Otolaryngology - Head and Neck Surgery	62	95	Surgery (66/420, Q1) Otorhinolaryngology (14/106, Q1)	1.505
The Lancet Respiratory Medicine	60	1,215	Pulmonary and Respiratory Medicine (1/131, Q1)	6.666

Source: Authors' elaboration based on Scopus database, June 2020.

According to the Scopus classification, documents can be classified into four different subject areas, namely: health sciences, life sciences, physical sciences and social sciences & humanities. However, these subject areas are strongly intertwined meaning that individual document can be classified in several subject areas at the same time. Accordingly, for the purposes of addressing the comprehensiveness of COVID-19 research, Figure 1 shows the Venn diagram of the presented subject areas and all the possible sets that can be made from them. This also makes it possible to determine

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 June 2020 doi:10.20944/preprints202006.0299.v1

the so-called pure sciences, covering only those documents belonging exclusively to just one subject area (without intersecting with other subject areas). According to the number of documents, health sciences contain a total of 8,896 documents of which 6,575 documents are identified as pure health sciences. Further, life sciences encompass a total of 2,549 documents of which 599 documents are considered as pure life sciences. Moreover, physical sciences include a total of 878 documents of which 314 documents belongs to pure physical science. Finally, social sciences & humanities cover 977 documents of which 323 are determined as pure social sciences & humanities. A comparison between different subject areas reveals that health sciences are the most relevant in COVID-19 research, while the second most relevant subject area is represented by life sciences. Moreover, physical sciences and social sciences & humanities seem to be the least popular so far. This is in line with the expectations. Namely, the first immediate response to COVID-19 pandemic is the protection of public health, while the real socio-economic consequences occur later. This path is also revealed by the recent scientific literature on COVID-19. Finally, some of the documents (n=173) are considered as multidisciplinary, making impossible to include them in the further bibliometric analysis.

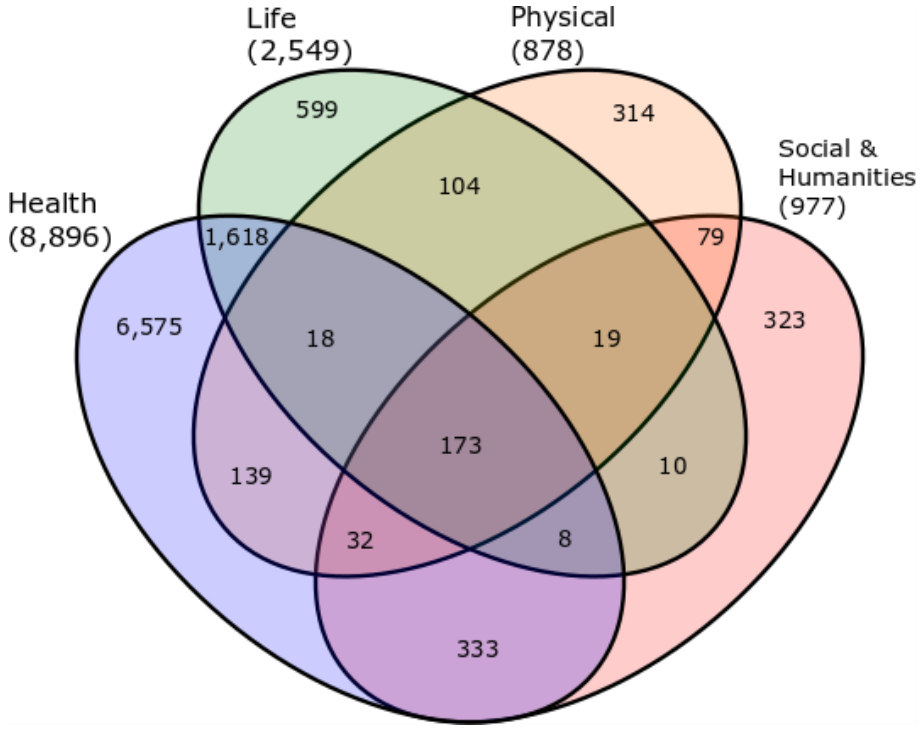

Figure 1. Venn diagram on COVID-19 research across different subject areas (January-May 2020). Source: Authors' elaboration based on Scopus database, June 2020.

Figure 2 presents the most relevant journals in COVID-19 research by subject area. It reveals the top 5 journals, providing the largest number of documents in health sciences, life sciences, physical sciences and social sciences & humanities separately. In health sciences, Journal of Medical Virology has the most documents (n=221), which is followed by The Lancet (n=196), the BMJ (n=187), JAMA – Journal of the American Medical Association (n=121) and Journal of Infection (n=121). For life sciences, due to strong interweaving with health sciences, the most relevant journal is Journal of Medical Virology, having the most documents (n=221), which is followed by Journal of Clinical Virology (n=70), Psychiatry Research (n=66), Brain, Behaviour and Immunity (n=64) ,and Eurosurveillance (n=50). In physical sciences, the most relevant journals are International Journal of Environmental Research (n=90), followed by Science of the Total Environment (n=88), International Journal of Advanced Science and Technology (n=52), Morbidity and Mortality Weekly Report (n=19) and Chaos, Solitons and Fractals (n=19). Finally, for social sciences the most relevant journals are Asian Journal of Psychiatry (n=60), followed by AIDS and Behavior (n=40), Economic and Political Weekly (n=36), Irish Journal of Psychological Medicine (n=27) and Social Anthropology (n=24).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 June 2020 doi:10.20944/preprints202006.0299.v1

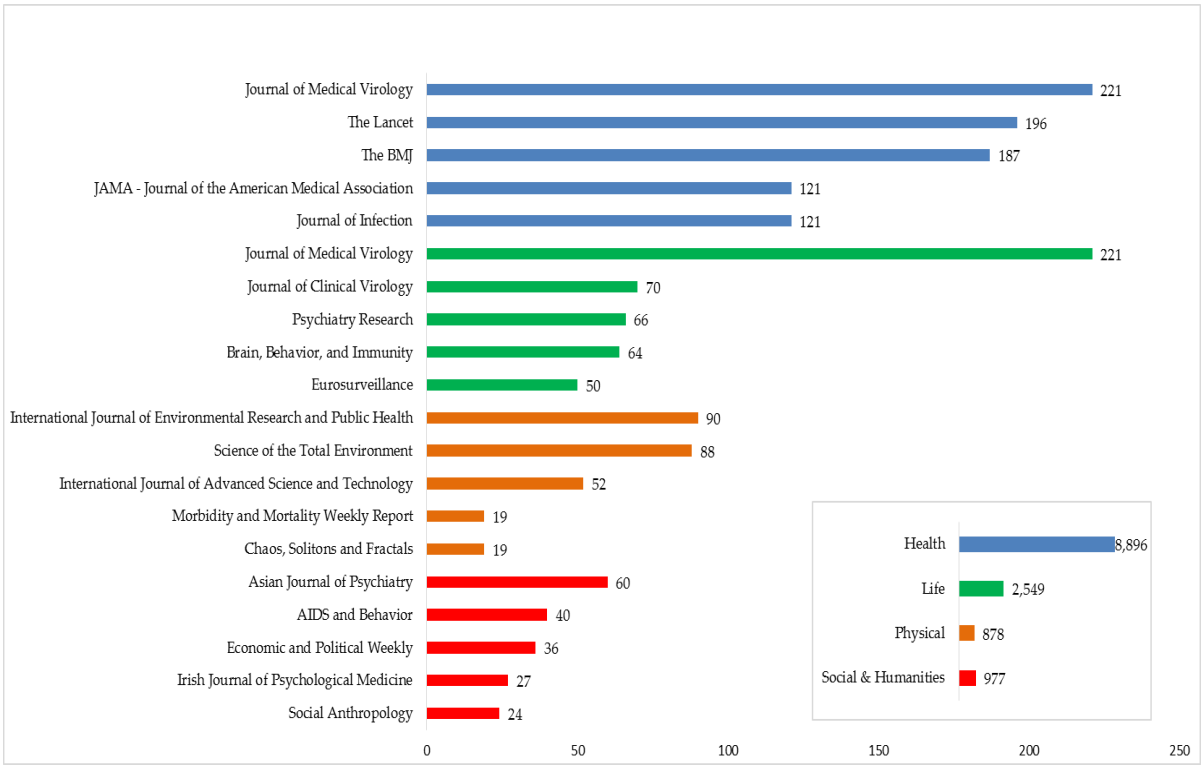

Figure 2. Most relevant journals in COVID-19 research by subject area (January-May 2020).
Source: Authors' elaboration based on Scopus database, June 2020.

Figure 3 presents the keyword co-occurrence network for (a) health sciences, (b) life sciences, (c) physical sciences and (d) social sciences & humanities separately. In order to ensure greater distinction between individual subject areas, only pure sciences (without intersecting with other sciences) are considered in the bibliometric analysis. Moreover, the bibliometric analysis is conducted on 100 most frequent (author and index) keywords by considering consolidation of the keywords describing the same phenomenon.

The bibliometric analysis reveals that research hotspots differ according to subject area. For health science, 3 clusters are identified, addressing the following topics: 1) pandemics; 2) risk factors and symptoms: and 3) mortality. Next, in the life science, 4 clusters are found, which are dealing with: 1) pandemics; 2) virology; 3) drug efficiency; and 4) vaccine. As regards physical science, 4 clusters are recognised, which are related to: 1) pandemics; 2) epidemiology; 3) viral disease and 4) air pollution. Finally, in social science, 13 clusters are identified, addressing the following topics: 1) pandemics; 2) epidemics; 3) viral disease in Asia; 5) globalization; 5) public health and economics; 6) social media; 7) tourism and education; 8) resilience; 9) technology; 10) financial markets; 11) crisis; 12) mental health and 13) elections.

The bibliometric analysis reveals that research hotspots differ according to subject area. For health science, 3 clusters are identified, addressing the following topics: 1) pandemics; 2) risk factors and symptoms: and 3) mortality. Next, in the life science, 4 clusters are found, which are dealing with: 1) pandemics; 2) virology; 3) drug efficiency; and 4) vaccine. As regards physical science, 4 clusters are recognised, which are related to: 1) pandemics; 2) epidemiology; 3) viral disease and 4) air pollution. Finally, in social science, 13 clusters are identified, addressing the following topics: 1) pandemics; 2) epidemics; 3) viral disease in Asia; 5) globalization; 5) public health and economics; 6) social media; 7) tourism and education; 8) resilience; 9) technology; 10) financial markets; 11) crisis; 12) mental health and 13) elections.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 June 2020 doi:10.20944/preprints202006.0299.v1

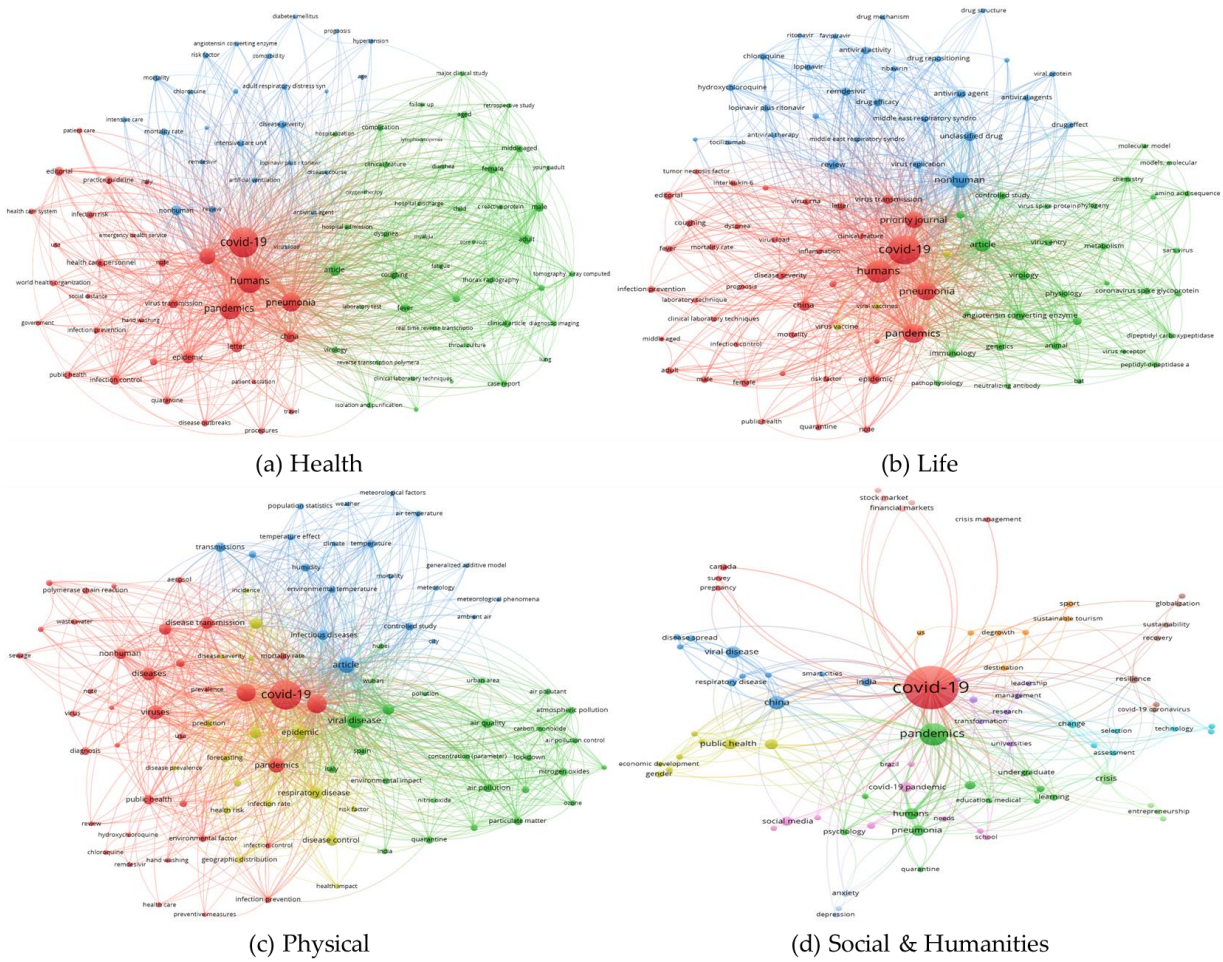

Figure 3. Keyword co-occurrence network in COVID-19 research by subject area (January-May 2020). Source: Authors' elaboration based on Scopus database, June 2020.

4. Conclusion

The outbreak of COVID-19 is a typical public health emergency, which due to its high infection rate makes it a huge threat not only to global public health but also to economic and social development. In order to be able to solve such kind of emergencies, it is necessary to fully understand the problem, its implications for different areas as well as the solutions that may be effective and efficient in addressing potential devastating consequences. Therefore, the scientific knowledge on COVID-19 is very important as it facilitates answering real-life questions. However, the extent of the current COVID-19 pandemic calls for in-depth knowledge allowing identification of numerous issues in different areas. Accordingly, this paper provides an extensive bibliometric analysis of COVID-19 research across science and social science research landscape by considering main subject areas and their relationships with one another.

The results show that a total of 10,344 documents related to COVID-19 were published in Scopus database in the first 5 months in 2020. They were written by 44,439 different authors, published in 1,978 different journals and together provide a total of 48,044 citations. Moreover, the most relevant journals in COVID-19 research cover about one-fifth of total documents and cover more than half of total citations. Most of them are ranked into the first quartile (Q1) and have a relatively high source normalized impact per paper (SNIP). Predominantly, they are subject to health sciences covering the subsubject areas of infectious diseases, general medicine and microbiology. When considering main subject areas separately, the results reveal that scientific disciplines are strongly intertwined, which calls for an in-depth analysis of individual subject area separately. According to the number of documents health science is the most relevant subject area in COVID-19 research, the second most

relevant subject area is life sciences, while physical sciences and social sciences & humanities seem to be the least popular. The results regarding journals reveal that Journal of Medical Virology is the most relevant journal for health science and life science, International Journal of Environmental Research for physical science and Asian Journal of Psychiatry for social science. Finally, the results of keyword co-occurrence analysis by main subject areas reveal different research hotspots for individual scientific disciplines, with a common point of pandemic. Health sciences are more focused on health consequences, while life sciences are more oriented towards drug efficiency. Moreover, physical sciences are more focused on environmental consequences, while social sciences are more oriented towards socio-economic consequences.

The paper highlights the importance of a comprehensive and in-depth approach considering different scientific disciplines in COVID-19 research. In order to address the health and socio-economic consequences of the current COVID-19 pandemic, COVID-19 must become the focus of the research in the near future. Namely, the understanding of the evolution of emerging scientific knowledge on COVID-19 is beneficial not only for scientific community but also for evidence-based policymaking in order to prevent and address the COVID-19 pandemic.

Author Contributions: A.A. supervised the work on the paper and revised it. D.R. wrote the paper. L.U. performed analysis. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: The authors acknowledge the financial support from the Slovenian Research Agency (research core funding No. P5-0093).

Conflicts of Interest: The authors declare no conflict of interest.

References

1 Fan, Y.; Zhao, K.; Shi, Z.L.; Zhou, P. Bat Coronaviruses in China. *Viruses* **2019**, *11*, 210, doi: https://doi.org/10.3390/v11030210.

2 Bogoch, I.I.; Watts, A.; Thomas-Bachli, A.; Huber, C.; Kraemer, M.U.; Khan, K. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. *J. Travel Med*. **2020**, *27*, 1-3, doi: https://doi.org/10.1093/jtm/taaa008.

3 Lin, Q.; Zhao, S.; Gao, D.; Lou, Y.; Yang, S.; Musa, S.S.; Wang, M.H.; Cai, Y.; Wang, W.; Yang, L.; He, D. A conceptual model for the outbreak of Coronavirus disease 2019 (COVID-19) in Wuhan, China with individual reaction and governmental action. *Int. J. Infect. Dis.* **2020**, *93*, 211-216, doi: https://doi.org/10.1016/j.ijid.2020.02.058.

4 Wu, J.T.; Leung, K.; Leung, G.M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* **2020**, *395*, 689-697, doi: https://doi.org/10.1016/S0140-6736(20)30260-9.

5 Gao, X; Yu, J. Public governance mechanism in the prevention and control of the COVID-19: information, decision-making and execution. *J. Chin. Gov.* **2020**, *5*, 178-197, doi: https://doi.org/10.1080/23812346.2020.1744922.

6 ECDC. COVID-19 situation update worldwide, as of 23 June 2020. Available online: https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases (accessed on 23 June 2020).

7 IMF. *World Economic Outlook, April 2020: The Great Lockdown*; IMF: Washington, USA, 2020.

8 OECD. *OECD Economic Outlook, June 2020*; OECD: Paris, France, 2020.

9 Hu, Y.; Chen, M.; Wang, Q.; Zhu, Y.; Wang, B.; Li, S.; Xu, Y.; Zhang, Y.; Liu, M.; Wang, Y.; Hu, Y. From SARS to COVID-19: A bibliometric study on emerging infectious diseases with natural language processing technologies. *Preprint* **2020**, doi: https://doi.org/10.21203/rs.3.rs-25354/v1.

10 Ram, S. Coronavirus Research Trends: A 50–Year Bibliometric Assessment. *Sci. Tech. Libr.* **2020**, *39*, 210-266, doi: https://doi.org/10.1080/0194262X.2020.1742270.

11 Tao, Z.; Zhou, S.; Yao, R.; Wen, K.; Da, W.; Meng, Y.; Yang, K.; Liu, H.; Tao, L. COVID-19 will stimulate a new coronavirus research breakthrough: a 20-year bibliometric analysis. *Ann. Transl. Med.* **2020**, 8, 528, doi: https://doi.org/10.21037/atm.2020.04.26.

12 Dehghanbanadaki, H.; Seif, F.; Vahidi, Y.; Razi, F.; Hashemi, E.; Khoshmirsafa, M.; Aazami, H. Bibliometric analysis of global scientific research on Coronavirus (COVID-19). *Med. J. Islam. Repub. Iran* **2020**, *34*, 354-362, doi: https://doi.org/10.34171/mjiri.34.51.

13 Hamidah, I.; Sriyono, S.; Hudha, M.N. A Bibliometric Analysis of Covid-19 Research using VOSviewer. *Indones. J. Sci. Technol*. **2020**, *5*, 34-41, doi: https://doi.org/10.17509/ijost.v5i2.24522.

14 Hossain, M.M. Current status of global research on novel coronavirus disease (Covid-19): A bibliometric analysis and knowledge mapping. *Preprint* **2020**, doi: https://doi.org/10.12688/f1000research.23690.1.

15 Lou, J.; Tian, S.J.; Niu, S.M.; Kang, X.Q.; Lian, H.X.; Zhang, L.X.; Zhang, J.J. Coronavirus disease 2019: a bibliometric analysis and review. *Eur. Rev. Med. Pharmacol. Sci.* **2020**, *24*, 3411-3421, doi: https://doi.org/10.26355/eurrev_202003_20712.

16 Nasab, F.R. Nasab, F.R., 2020. Bibliometric Analysis of Global Scientific Research on SARSCoV-2 (COVID-19). *Preprint* **2020**, doi: https://doi.org/10.1101/2020.03.19.20038752.

17 Wang, C.; Lim, M.K.; Zhao, L.; Tseng, M.L.; Chien, C.F.; Lev, B. The evolution of Omega-The International Journal of Management Science over the past 40 years: A bibliometric overview. *Omega* **2020**, *93*, 102098, doi: https://doi.org/10.1016/j.omega.2019.08.005.