



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



18th International Learning & Technology Conference 2021

A Bibliometric Analysis of COVID-19 Vaccines and Sentiment Analysis

Akila Sarirete¹

*Computer Science Department
College of Engineering, Effat University
Energy and Technology Research Center, Effat University
Jeddah, Saudi Arabia*

Abstract

Recent statistical and social studies have shown that social media platforms such as Instagram, Facebook, and Twitter contain valuable data that influence human behaviors. This data can be used to track, fight, and control the spread of the COVID-19 and are an excellent asset for analyzing and understanding people's sentiments. Current levels of willingness to receive a COVID-19 vaccination are still insufficient to achieve immunity standards as stipulated by the World Health Organization (WHO). The present study employs bibliometric analysis to uncover trends and research into sentiment analysis and COVID-19 vaccination. A range of analyses is conducted using the open-source tool VOSviewer and Scopus database from 2020-2021 to acquire a deeper insight and evaluate current research trends on COVID-19 vaccines. The quantitative methodology used generates various bibliometric network visualizations and trends as a function of publication metrics such as citation, geographical attributes, journal publications, and research institutions. Results of network visualization revealed that understanding the the-state-of-the-art in applying sentiment analysis to the COVID-19 pandemic is crucial to local government health agencies and healthcare providers to help in neutralizing the infodemic and improve vaccine acceptance.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 18th International Learning & Technology Conference 2021

Keywords: COVID-19; Vaccine, Sentiment Analysis, Bibliometric, VOSviewer.

1. Introduction and Background

Coronavirus (COVID-19) is a contagious respiratory disease initially reported in late December 2019 in Wuhan, China, and has since spread rapidly throughout the world. COVID-19 virus causes acute respiratory diseases starting from the common cold to deadly pneumonia of various degrees. According to the World Health Organization (WHO) [1], the COVID-19 infects people of all ages; however, research indicates that two categories of people are more likely to have severe COVID-19 disease, the elderly people, and those with pre-existing medical issues. On

¹Corresponding author. *E-mail address:* asarirete@effatuniversity.edu.sa

March 11, 2020, the WHO declared the new coronavirus (COVID-19) outbreak a pandemic and urged for the need for a worldwide commitment to fight and control the pandemic [2] [3]. The WHO created a comprehensive guideline strategy to slow down the pandemic while looking for efficient and safe vaccines. The guideline is based on measures such as: public health measures, physical and social measures, restriction of mass gatherings, closing of schools, universities, shopping malls, and restaurants, and restricting travels to places where the COVID-19 is spreading out. While underling that to protect others, everyone must protect himself from COVID-19. This situation created a severe challenge for all countries to monitor and slow down the pandemic using different measures while avoiding the 3Cs (closed spaces, crowded, or closed contact). One year later, as of March 11, 2021, COVID-19 continues to have a substantial impact on the everyday life of people and has spread to more than 210 countries and territories, and affected almost 120 million people worldwide with more than 2.7 million recorded deaths [4].

New strategies, physical and social measures have been adopted to accelerate the fight against the COVID-19, including the development of vaccines [5]. The top priority for 2021 is to keep battling COVID-19, restore and improve current health systems, accelerate access to COVID-19 treatment, and provide equitable and safe vaccines for all. WHO emphasized that vaccine access should be fair and honest for every country to protect its people. However, as COVID-19 is still a controversial global topic on social media with mixed emotions, positive and negative, due to the infodemic, the WHO also took proactive steps in immunizing the public against misinformation and help in neutralizing the infodemic that is undermining the outbreak response [6].

For a vaccine to be successful, many people need to get vaccinated and prevent the virus from spreading out. Understanding the state-of-the-art development of COVID-19 vaccines and the corresponding public sentiment regarding vaccination is crucial in fighting this pandemic and planning future research efforts to monitor and identify any genetic changes in the virus that allow us to understand the potential impact of mutations. In recent months, there has been a considerable amount of studies towards vaccines against the COVID-19 virus. Multiple COVID-19 vaccines have already been approved in many countries; these vaccines include Pfizer–BioNTech, Moderna, Sputnik V, Sinovac, and Oxford–AstraZeneca [7], [8]. According to Statista [9], as of February 12, 2021, nearly three billion AstraZeneca/Oxford's vaccines, one billion Pfizer-BioNTech vaccines, and more than six hundred million Moderna vaccines were pre-purchased. Figure 1 shows the share of the total population who received at least one vaccine dose [10].

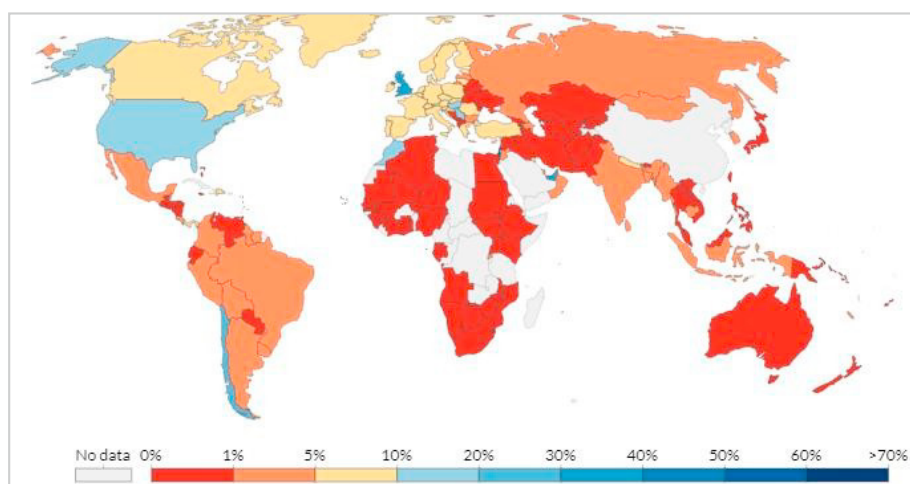


Fig. 1: Share the total population that received at least one vaccine dose as of March 11, 2021, [10].

Wang et al. [11] indicated that the United States, China, Italy, and the United Kingdom contributed nearly 60% of the WOS publications on global scientific research on COVID-19. Almost 80% of them are published in top journals. Yet, there is an urgent need for international cooperation in scientific research to fight the COVID-19. A comprehensive review of the global efforts on developing COVID-19 vaccines has been conducted by Yingzhu et al. [7] using CAS content collection (a division of the American Chemical Society). The report examined the advantages and disadvantages of COVID-19 vaccines and future research activity trends. The study concludes by stressing the many challenges and unanswered questions in the development of the COVID-19 vaccine. Yet, the study offered an essential step towards understanding the future trends of the vaccine for a safe and effective vaccine. Other studies conducted an in-depth investigation into the various types of vaccination, their compositions, benefits, safety, efficacy, and potential limits [11] [12] [13]. A comprehensive study with detailed information on COVID-19 vaccines was conducted by Ahmed et al. [14] using the Web of Science database (WOS) with 916 documents in 376 journals. In another study, Zhang et al. [15] used the WOS database to investigate the 100 top-cited articles on vaccines using bibliometric analysis

and included citation time, density, journal name and impact factor, article type, and category country of origin. In a survey of people from 19 countries conducted in June 2020 and published in February 2021 by Nature Medicine [16], 71.5% of participants said they would be very or somewhat likely to get a COVID-19 vaccine, while 48.1% said they would accept their employer's recommendation. In this study, the acceptance rates varied from nearly 90% in China to less than 55% in Russia and are strongly related to the level of trust in government information. This study indicates that the current willingness to receive a COVID-19 vaccination is insufficient to achieve community immunity standards, as noted in the WHO. In another survey conducted on December 2020 by KFF (Kaiser Family Foundation) [17], 71% of respondents said they would agree to take the COVID-19 vaccine if it is free of charge and safe, while about 27% remained hesitant, fearing side effects. It is essential to understand and overcome the wide variation of willingness to accept a COVID-19 vaccine from these studies. This variation could potentially delay global control of the pandemic and the subsequent societal and economic recovery.

In the light of this motivation, this paper uses bibliometric analysis to identify trends and research in COVID-19 vaccines based on sentiment analysis. A range of analyses is conducted based on Scopus articles to acquire a deeper insight and ongoing research in COVID-19 vaccines. The remainder of the paper is as follows. Section 2 describes the research methodology used to analyze COVID-19 vaccines and sentiment analysis; section 3 presents and discusses the present results, and section 4 highlights the study's main conclusions.

2. Research Methodology to Capture Trends in COVID-19 Vaccines and Sentiment Analysis

While managing the pandemic using social distancing and physical measures, recent statistical and social studies have shown that social media platforms such as Instagram, Facebook, and Twitter contain many valuable insights that can be used to manage the pandemic and its impacts. These platforms would help align human behavior with the recommendations from the WHO, local health organizations, and public health experts. According to research publications [18], many pandemics would have been quickly monitored if public health authorities and experts in the field had incorporated social media as one of the main components in helping fight the pandemic. Sentiment analysis, also known as opinion mining, is the process of studying and detecting people's attitudes and feelings through text analysis via social media. This technique of social media monitoring has been applied in business, stock market, education, tourism and is currently involved in healthcare [19] [20]. The quantitative methodology used in this study to generate various bibliometric visualizations and trends is based on retrieving a set of publications from the Scopus database. First, publications related to COVID-19 vaccines and sentiment analysis were retrieved and collected using the Scopus database. Second, the retrieved data were exported into Microsoft Excel for Windows. The generated CSV file is then imported by VOSviewer software version 1.6.16 [21] to construct bibliometric networks in the third stage. The bibliometric analysis extracts quantitative information about the publication metrics such as citation and co-citation, geographical attributes, author co-authorship, journal, and research institution. VOSviewer additionally offers Text Mining mapping to build and visualize coexisting networks of key terms extracted from the scientific literature. The analysis was used to determine the most and current state of COVID-19 vaccines and sentiment analysis.

3. Results and discussion

This study used the Scopus collection database for documents published in 2020/2021 and limited to English. The initial literature on COVID-19 vaccines documents has been scanned using the search string TITLE-ABS-KEY (((("covid19" OR "covid-19" or "SARSCov2" or "SARS-CoV-2" or "mRNA" or "2019-nCoV" or "corona" or "corona virus" or "Coronavirus" or "Novel-Coronavirus") AND ("pfizer" or "Pfizer- Biotech" or "Moderna" or "AstraZeneca" or "AstraZeneca-Oxford" or "Johnson & Johnson" or janssen or "Sputnik" or "Sinopharm" or "Sinova" or "vaccin*")))).

The initial search, which includes Title, Abstract, and Keyword, resulted in 8083 documents. The author refined the retrieved documents using subject area, document type, publication stage, source type, and Title only, resulting in 2109 documents published in more than 160 journals. The retrieved documents consist of 1567 articles, 493 reviews, 41 short surveys, three book chapters, three conference papers, and 2 data papers.

In 2021, the number of documents published is more than double (1446 documents) compared to 2020 (663 documents). More than 100 countries produced at least 1 document, 63 countries had more than five documents, 49 countries producing at least ten documents, 30 countries producing at least 20 documents, 16 countries producing at least 40 documents, and seven countries produced at least 100 documents, namely the United States, the United Kingdom, China, India, Italy, Canada, and Germany.

One hundred sixty active institutions published at least five documents on the COVID-19 and vaccines, 91 institutions published at least ten documents, and 19 institutions published at least 20 documents. The top-5 subject areas on COVID-19 and Vaccines were *Medicine* (49%), *Immunology and Microbiology* (19%), *Pharmacology, Toxicology*

and *Pharmaceutics* (15%), *Biochemistry, Genetics and Molecular Biology* (9%), and *Multidisciplinary* (4%). The top-5 journals with the maximum number of published documents were *Vaccines* (14%), *BMJ Clinical Research* (10%), *Human Vaccines and Immunotherapeutic* (3%), *New England Journal of Medicine* (3%), and *Frontiers in Immunology* (2%). Table 1 shows the top-10 active affiliations with 2.44 % from Harvard Medical School, followed by the University of Oxford with 2.25%, and the University of Pennsylvania with 2.10%.

Table 1 Top-10 active affiliations.

Ranking	Affiliation	Record	%
1	Harvard Medical School	50	2.44%
2	University of Oxford	46	2.25%
3	University of Pennsylvania	43	2.10%
4	National Institutes of Health NIH	42	2.05%
5	London School of Hygiene & Tropical Medicine	35	1.71%
6	Harvard T.H. Chan School of Public Health	30	1.46%
7	Imperial College London	29	1.42%
8	The University of North Carolina at Chapel Hill	28	1.37%
9	University of Pennsylvania Perelman School of Medicine	27	1.32%
10	Emory University	25	1.22%

The author used VOSviewer software (version 1.6.16) to perform the analysis and counting method for co-authorship, co-occurrence, citation, and co-citation for authors, organizations, and countries [21]. With this tool, the author created and visualized maps based on a network of Scopus scientific publications, authors, research organizations, countries, and keywords. To construct a network, the retrieved data from the Scopus database were exported as a CSV file and imported by VOSviewer, which provides a network visualization, an overlay visualization, and a density visualization. Labels represent items in the network visualization using circles. The weight of an item determines the size of the label and the circle around it. The label for some items may not be visible to prevent possible overlapping labels. An item's color is determined by the cluster to which it belongs. The lines between items represent Links. In the present study, publications co-authored by many countries were excluded, and the maximum number of countries per publication was set to 25. The minimum number of publications per country was fixed at 5, with a minimum number of citations for a country set to 5. Out of 188 countries, only 59 met the thresholds. The USA was the leading country (documents = 658, citations = 13390, total link strength (TLS) = 531), followed by United Kingdom (documents = 216, citations = 4905, TLS = 357), China (documents = 165, citations = 3555, TLS = 144), India (documents = 146, citations = 1170, TLS = 103), and Italy (documents = 131, citations = 769, TLS = 146). Figure 2 shows the co-authorship country visualization network map. Accordingly, five clusters were formed: cluster 1 indicated by a red color includes 20 countries led by India with 146 documents; cluster 2 indicated by a green color includes 16 countries led by Italy with 131 documents; cluster 3 indicated by a blue color includes ten countries led by the United States with 658 documents; cluster 4 indicated by a yellow color includes four countries led by China with 165 documents; and cluster 5 indicated by a purple color includes four countries led by South Africa with 31 documents.

For the co-occurrence author keywords, the minimum number of occurrences of a keyword was set at 5. Out of a total of 8256 keywords, only 1429 keywords met the threshold. The top-5 dominant keywords occurrence were COVID-19 (1196, TLS = 32400), Coronavirus disease (915, TLS = 33563), Sars-cov-2 (862, TLS = 26137), Vaccination (823, TLS = 2521), and COVID-19 Vaccines (634, TLS = 19291). The visualization network map of the

co-occurrence of all keywords is shown in Figure 3 and includes 5 clusters (red, green, blue, yellow, and purple colors.) with three dominant ones, COVID-19, immunology, and virus vaccine.

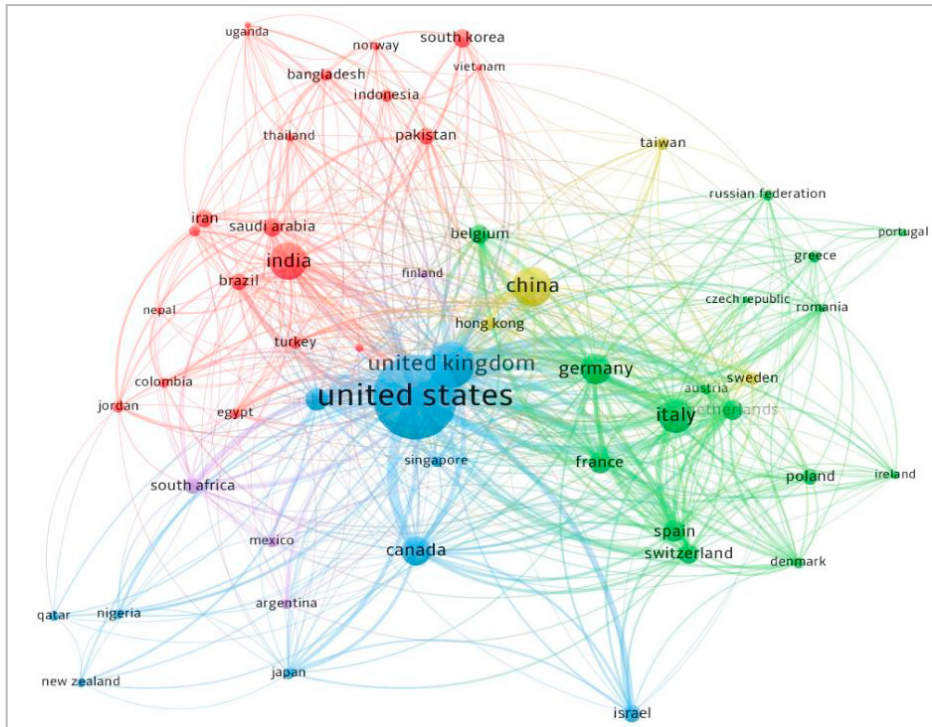


Fig. 2. VOSviewer Co-authorship country visualization network map.

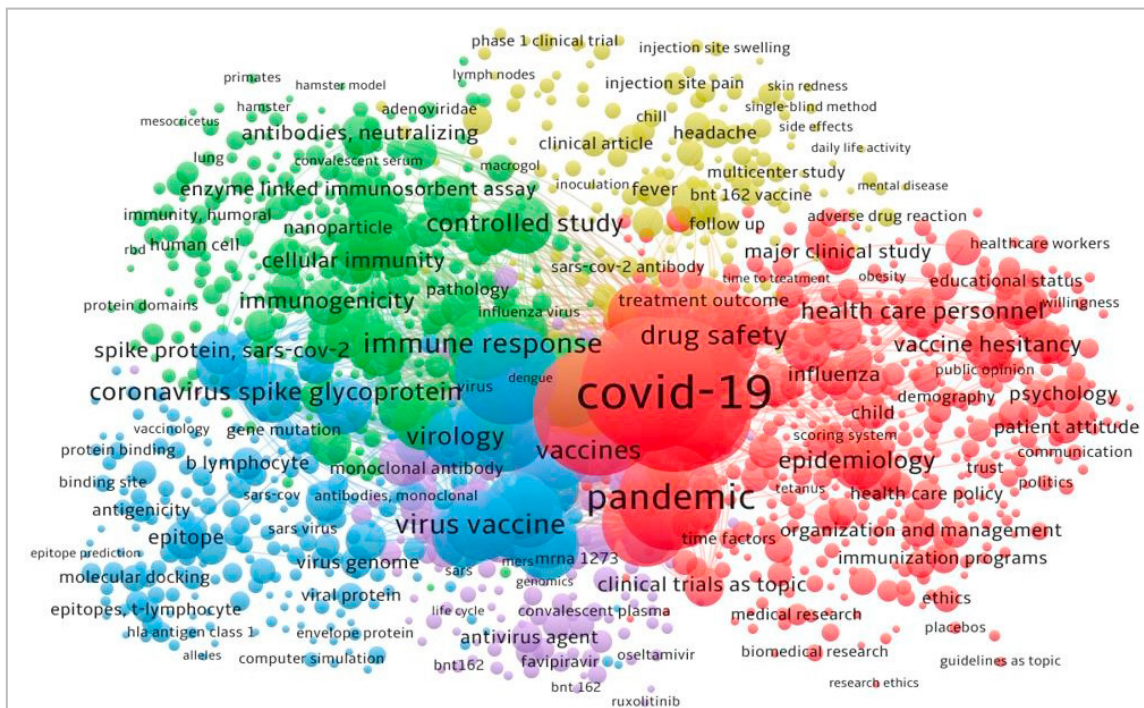


Fig. 3. VOSviewer network visualization map for author keywords co-occurrence, COVID-19, and vaccine.

For the sentiment analysis, we added strings for sentiment analysis to the original string result represented by figures 2 and 3. The author used frequent keywords such as ‘sentiment analysis’, ‘social media’, ‘text mining’, ‘Twitter’, etc. The string added is ("sentiment analysis" or "social media" or "opinion mining" or "text mining" or "Twitter" or "social media analytics" or "Big data" or "machine learning" or "sentiment classification" or "natural language processing" or "information retrieval" or "text classification" or "web mining" or "social networks" or "NLP" or "polarity classification" or "feature selection" or "Opinion extraction" or "ontology" or "unsupervised learning" or "Twitter" or "Facebook") and (limit-to (language, "English")) and(limit-to (pubyear , 2021) or limit- to (pubyear , 2020)). A total of 1237 documents, distributed as shown in figure 4, were retrieved from the Scopus dataset.

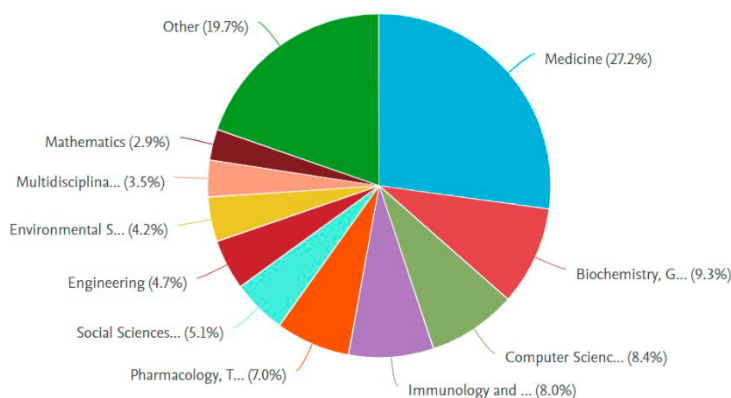


Fig. 4. Distribution of the retrieved documents by document area.

Using VOSviewer, 3656 keywords were retrieved. The minimum number of occurrences of a keyword was set at 5. 250 keywords met the threshold out of 3656 keywords. The top-5 dominant keywords occurrence were COVID-19 (344, TLS = 2923), Sentiment Analysis (241, TLS = 1733), Social Media (199, TLS = 2287), Twitter (99, TLS = 969), and Coronavirus (76, TLS = 836). Liu [22] debated whether to call the field sentiment analysis or opinion mining. Our results from the Scopus database indicated that sentiment analysis had been used 241 times while opinion mining was used 15 times only, as shown in figure 5, along with other keywords retrieved from the Scopus database. We included the infodemic as an important step in immunizing the public against misinformation undermining the outbreak response [6]. The visualization network map of the co-occurrence of all keywords is shown in figure 6. It includes 8 clusters (red, green, blue, yellow, and purple colors.) with three dominant ones, COVID-19, Sentiment Analysis, and Social Media. The top 10 active countries of research on COVID-19 and sentiment analysis are shown in figure 7.

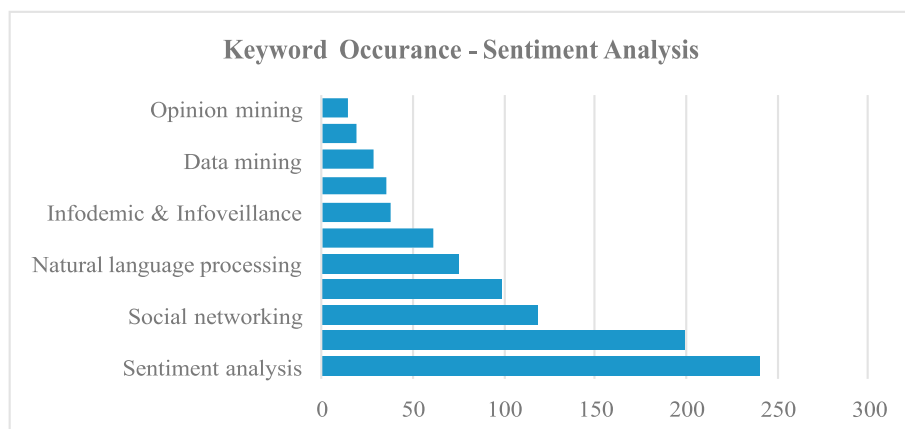


Fig. 5. Distribution of the keywords used in COVID-19 and Sentiment Analysis

of the COVID-19 pandemic and people's sentiments are crucial in planning future efforts towards fighting COVID-19 and improving vaccine acceptance.

Acknowledgments

The author gratefully acknowledges the support of Effat College of Engineering at Effat University, Jeddah, Saudi Arabia.

References

- [1] "Situation Report - 51." <https://www.who.int/publications/m/item/situation-report---51> (accessed February 28, 2021).
- [2] *WHO Director-General's opening remarks at the media briefing on COVID-19 - March 11 2020*. Accessed: February 20, 2021. [Online]. Available: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
- [3] D. Cucinotta and M. Vanelli, "WHO Declares COVID-19 a Pandemic," *Acta Bio-Medica Atenei Parm.*, vol. 91, no. 1, pp. 157–160, Mar. 2020, doi: 10.23750/abm.v91i1.9397.
- [4] "Coronavirus Graphs: Worldwide Cases and Deaths - Worldometer." <https://www.worldometers.info/coronavirus/worldwide-graphs/> (accessed March 10, 2021).
- [5] N. Lurie, M. Saville, R. Hatchett, and J. Halton, "Developing COVID-19 Vaccines at Pandemic Speed," *N. Engl. J. Med.*, vol. 382, no. 21, pp. 1969–1973, May 2020, doi: 10.1056/NEJMp2005630.
- [6] "Immunizing the public against misinformation." <https://www.who.int/news-room/feature-stories/detail/immunizing-the-public-against-misinformation> (accessed March 11, 2021).
- [7] "COVID-19 vaccine tracker and landscape." <https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines> (accessed March 11, 2021).
- [8] "COVID-19: List of countries which have already approved emergency use of vaccine," *Hindustan Times*, December 12, 2020. <https://www.hindustantimes.com/world-news/covid-19-list-of-countries-which-have-already-approved-emergency-use-of-vaccine/story-pflubv6xjzCQXR9K7bJ4M.html> (accessed February 13, 2021).
- [9] "COVID-19 vaccines ordered by manufacturer 2021," *Statista*. <https://www.statista.com/statistics/1195971/number-of-covid-19-vaccine-doses-by-manufacturer/> (accessed February 21, 2021).
- [10] "Share of people who received at least one dose of COVID-19 vaccine," *Our World in Data*. <https://ourworldindata.org/grapher/share-people-vaccinated-covid> (accessed February 21, 2021).
- [11] J. Wang, Y. Peng, H. Xu, Z. Cui, and R. O. Williams, "The COVID-19 Vaccine Race: Challenges and Opportunities in Vaccine Formulation," *AAPS PharmSciTech*, vol. 21, no. 6, p. 225, Aug. 2020, doi: 10.1208/s12249-020-01744-7.
- [12] O. Sharma, A. A. Sultan, H. Ding, and C. R. Triggler, "A Review of the Progress and Challenges of Developing a Vaccine for COVID-19," *Front. Immunol.*, vol. 11, 2020, doi: 10.3389/fimmu.2020.585354.
- [13] "COVID 19 Vaccination Efficacy And Safety Literature Review | MaplesPub." <https://maplepub.com/article/COVID-19-Vaccination-Efficacy-and-Safety-Literature-Review> (accessed February 8, 2021).
- [14] T. Ahmad, M. A. Murad, M. Baig, and J. Hui, "Research trends in COVID-19 vaccine: a bibliometric analysis," *Hum. Vaccines Immunother.*, pp. 1–6, Mar. 2021, doi: 10.1080/21645515.2021.1886806.
- [15] Y. Zhang, L. Quan, B. Xiao, and L. Du, "The 100 top-cited studies on vaccine: a bibliometric analysis," *Hum. Vaccines Immunother.*, vol. 15, no. 12, pp. 3024–3031, Dec. 2019, doi: 10.1080/21645515.2019.1614398.
- [16] J. V. Lazarus et al., "A global survey of potential acceptance of a COVID-19 vaccine," *Nat. Med.*, vol. 27, no. 2, pp. 225–228, Feb. 2021, doi: 10.1038/s41591-020-1124-9.
- [17] C. Muñana and 2020, "KFF COVID-19 Vaccine Monitor: December 2020," *KFF*, December 15, 2020. <https://www.kff.org/coronavirus-covid-19/report/kff-covid-19-vaccine-monitor-december-2020/> (accessed February 14, 2021).
- [18] A. H. Alamoodi et al., "Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review," *Expert Syst. Appl.*, vol. 167, p. 114155, Apr. 2021, doi: 10.1016/j.eswa.2020.114155.
- [19] L. Abualigah, H. Alfar, M. Shehab, and A. Hussein, "Sentiment Analysis in Healthcare: A Brief Review," 2020, pp. 129–141. doi: 10.1007/978-3-030-34614-0_7.
- [20] K. Cortis and B. Davis, "Over a decade of social opinion mining: a systematic review," *Artif. Intell. Rev.*, Jun. 2021, doi: 10.1007/s10462-021-10030-2.
- [21] "VOSviewer - Visualizing scientific landscapes," *VOSviewer*. <https://www.vosviewer.com/> (accessed February 9, 2021).
- [22] L. Zhang and B. Liu, "Sentiment Analysis and Opinion Mining," in *Encyclopedia of Machine Learning and Data Mining*, C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2017, pp. 1152–1161. doi: 10.1007/978-1-4899-7687-1_907.