

Original Research

Thematic Evolution and Dissemination of Scholarly Outputs During Health Crises: A Scientometric and Altmetric Analysis of COVID-19 Research

Vahideh Zarea Gavvani

Professor, Department of Medical Library and Information Science, School of Management and Medical Informatics, Tabriz Health Services Management Research Center, Tabriz University of Medical Sciences, Tabriz, Iran.
Zarehv@tbzmed.ac.ir / vgavvani@gmail.com

ORCID iD:

<https://orcid.org/0000-0002-3278-6859>

Mina Mahami-Oskouei

Assistant Prof., Department of Medical Library and Information Science, School of Management and Medical Informatics, Tabriz University of Medical Sciences, Tabriz, Iran.

Corresponding Author: mahami.m@tbzmed.ac.ir / mmahami24@yahoo.com

ORCID iD:

<https://orcid.org/0000-0001-8561-8042>

Received: 06 December 2024

Reviewed: 23 December 2024

Accepted: 09 July 2025

Abstract

The COVID-19 pandemic represented a critical case for studying the rapid dynamics of research and its public dissemination during a global crisis. This study addresses the need for a comprehensive understanding of these changes by examining the thematic evolution of COVID-19 research and its social attention during the initial phase of the pandemic (2020–2021). A combined scientometric and Altmetric approach was utilized. Data, encompassing 16,422 articles indexed in PubMed and 1,751,396 social mentions (primarily from Twitter and news media) extracted via Altmetric.com, were analyzed. Keyword co-occurrence network analysis (using VOSviewer) and descriptive statistics were employed to identify thematic clusters, track their evolution, and measure the social reach of the publications. Findings revealed a significant shift in research priorities. Initially, the focus was concentrated on virology, clinical sciences, and public health. As the pandemic progressed, the thematic landscape broadened dramatically to encompass psychological, social, economic, and policy-related topics, signaling a shift from immediate biological concerns to broader societal impacts. Twitter emerged as the dominant dissemination channel, hosting the vast majority of social media mentions, underscoring its crucial role in health communication during the crisis. The network analysis confirmed these thematic shifts, highlighting the growing centrality of social science subjects over time. In conclusion, this thematic analysis serves as a predictive tool, illustrating how public health crises rapidly redefine research agendas, shifting them from fundamental medical science to interdisciplinary social issues. Policymakers and research managers can leverage these insights to establish interdisciplinary research strategies and enhance preparedness for future crises by anticipating the evolving socio-scientific needs.

Keywords: Altmetrics, Health Information Dissemination, COVID-19 Pandemic, Bibliometrics, Thematic Analysis.

Introduction

The COVID-19 pandemic was not only an unprecedented public health crisis but rapidly became a watershed moment in the history of scientific production and communication. Amidst scientific uncertainty and a global surge of misinformation (or "infodemic"), the urgent need to generate credible evidence and disseminate it rapidly through both scholarly journals and social media became a critical imperative (Zarocostas, 2020). In such crises, these conditions can both accelerate research growth and establish a fragile relationship between researchers and the public; on the one hand, scientific output increases at an unparalleled pace, and on the other, public trust and the legitimacy of research may be pressured (Ahmed & Mao, 2022).

One study demonstrated how the prioritization and acceptance of research proposals can shift significantly during crises, underscoring the fragile relationship between researchers and society (Riley, Xie & Khurshid, 2021). Conversely, prior research has shown that at the onset of crises, such as pandemics, initial research is often fragmented and lacks cohesion, making it challenging for academics to identify the core research trajectories (Bundy, Pfarrer, Short & Coombs, 2017).

In the context of COVID-19, the initial wave of publications primarily focused on the immediate clinical and biological aspects of the disease. However, new themes gradually entered the scientific literature, including mental health, virtual education, health policy, and socio-economic dimensions (Furstenau et al., 2021). This rapid thematic transformation underscores the crucial importance of analyzing thematic evolution, as only by examining these conceptual shifts can we fully understand how the scientific community redefines its trajectory in response to a crisis.

In this context, tools like Altmetrics play a pivotal role in revealing patterns of public and scientific attention. By tracking feedback and engagement with scholarly articles in the public sphere, Altmetrics allows for the real-time identification of "hot" and highly debated themes (Vosoughi, Roy & Aral, 2018; Elmore, 2018). The integration of Altmetric and bibliometric analyses enables researchers and policymakers to grasp both the shifts in scientific priorities and their correlation with salient socio-economic issues (Mardiani, Riniati, Mokodenseho, Podomi & Mamonto, 2024).

Despite the explosive growth of publications during the COVID-19 crisis, significant gaps remain. Most previous studies have either focused on very short timeframes or examined only partial subjects, such as vaccines or economic dimensions, failing to provide a comprehensive picture of the initial wave of thematic evolution (Haghani, Bliemer, Goerlandt & Li, 2020; Ioannidis, Cripps & Tanner, 2022). Thematic analysis in scientific research allows policymakers and researchers to understand the emergence and significance of novel issues, helping them to improve their responses and preparedness for future crises (Ozuem, Willis & Howell, 2022). Although COVID-19 is a unique and prominent case in contemporary history, other health crises may follow similar patterns.

Therefore, focusing on the period from 2020 to 2021 is essential, as this timeframe represents the initial formation and stabilization of COVID-19 research patterns. During this period, the scientific community produced knowledge at an unprecedented speed, and fundamental changes occurred in the structure and content of research. Examining this era can provide a clear understanding of the scientific reaction to a global crisis and the primary pathways of thematic transformation during the initial years of the pandemic.

Accordingly, the present study employs a scientometric approach, using co-occurrence

network analysis, to examine the thematic evolution of COVID-19 research between 2020 and 2021. The aim is twofold: to contribute to the theoretical enrichment of crisis scientometrics and to provide practical insights for health policy-making and future crisis preparedness by identifying dominant themes, emerging topics, and conceptual developments.

Materials and Methods

The current study was a cross-sectional, analytical study using large-scale data from Altmetric.com. It examined the relationship between altmetrics data and worldwide confirmed COVID-19 cases. It used PubMed articles on novel Coronavirus (COVID-19) to track the aggregation of attention to these articles on Twitter, Facebook, and in the news media via Altmetric.com. Moreover, to assess the involvement of different research subjects during the pandemic, a network of subject keyword co-occurrence was analyzed. The entire study process, including the search, screening, collection of Altmetric data, and network analysis, is schematically illustrated in Figure 1.

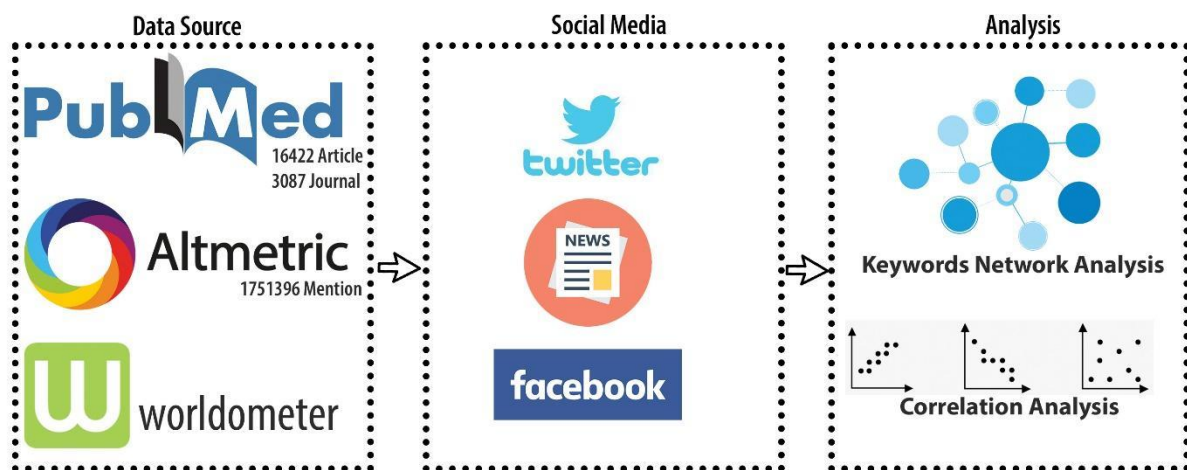


Figure 1: The overall scheme of the research

Data acquisition

The metadata for articles were extracted from a reproducible search in <https://www.ncbi.nlm.nih.gov/pubmed/>. The following search terms and strategy were employed to retrieve information from PubMed and Altmetric.com. ("Novel coronavirus 19" [text word] OR "novel coronavirus 2019" [text word] OR "2019-nCoV" OR "coronavirus 2019" [text word] OR "Covid 19"[text word] OR "Covid-19" OR "Covid 2019"[text word] OR "nCov 2019"[text word])

Keyword validation

Coronavirus refers to a large family of viruses that usually cause respiratory diseases. Common types of Coronavirus that have infected humans in recent years are known as SARS and MERS, also called coronavirus 18. Therefore, searching for "Coronavirus" retrieves a considerable number of articles on other types of Coronaviruses. The confounding keywords in the search strategy were controlled by the search-review strategy. Once the search was performed for the first time, the title and abstract of the retrieved articles were reviewed to evaluate the relevance and validation. The confounding keywords were identified and excluded

from the search, and the search was re-run with the eligible terms. The results were reviewed once again to reach a valid list of keywords. To perform a comprehensive search strategy, we used the [Text Word] tag. All terms from one set were combined with the 'OR' Boolean Operator, and 'AND' was used to combine all keyword sets in one search string. To ensure data validity and reproducible results, the retrieved set of articles from PubMed was reviewed by two independent researchers, and keyword refinement was performed in two stages to prevent the inclusion of irrelevant articles. Additionally, the consistency of the Altmetric data with the retrieved articles was checked, and only valid data were included in the analyses.

Inclusion and exclusion criteria

All articles published within the specified timeframe, including original research, reviews, short reports, and editorials, were included in the study. Articles lacking a Digital Object Identifier (DOI) or those with unverified bibliographic information were excluded. Furthermore, in the Altmetric data, only posts with an identified IP address were considered. Anonymous or bot-generated posts were filtered out as they could introduce noise and potentially compromise the validity of the results. It is important to note that the exclusion of this category of data accounted for less than 5% of the total volume of mentions and therefore had no significant impact on the findings.

The data collection period was limited to January 1st, 2020, to December 1st, 2021. This timeframe was chosen because it captures the initial phase of the pandemic and the emergence of the first wave of scientific research, reflecting the period of most rapid change in the COVID-19 scientific literature.

Altmetric data collection

Altmetric data were extracted via the official Altmetric.com API and received in CSV format. This dataset encompassed mentions on social networks (Twitter and Facebook), news outlets, and blogs. Microsoft Excel was used for data control and cleaning, after which the data were transferred to network analysis software.

Data analysis

Initially, descriptive statistics were calculated to examine the coverage and frequency of mentions. Subsequently, keyword co-occurrence network analysis was employed to analyze the thematic evolution. The software used for this section was VOSviewer (version 1.6.17). The unit of analysis was defined as the keywords tagged in the Altmetric database, and the entry threshold for these keywords into the network was set at a minimum of 15 occurrences. The association strength method was used to calculate link weights. Articles in the altmetric database were categorized based on 17 subject categories. To assess the significance of the themes and their relationships, network metrics, including degree, betweenness centrality, and density, were also computed. The subjects with larger circles have a greater impact in forming relationships with other subjects. On the other word these subjects (with bigger circles) are middle nodes that link other nodes to each other. Finally, thematic networks were plotted and compared for different time intervals to systematically and clearly identify changes in dominant themes, emerging topics, and evolutionary paths.

Results

We identified 16,422 documents from the PubMed search, which were published in 3087 journals and had at least one mention in altmetric database. Tracking articles on Altmetric.com led to the identification of 1,751,396 mentions across various metrics. Table 1 summarizes the mentions obtained from altmetric.com. We analyzed only three metrics, including Twitter, Facebook, and the news media.

Table 1

Frequency of distribution of mentions in alternative metrics

Category	Subcategory	Mentions	Percentage of total (%)
Social Media ~97%	Twitter	1,695,915	96.83
	Facebook	2,449	0.14
	Reddit posts	1,735	0.10
News and Blogs ~3%	News	47,145	2.69
	Blogs	3,019	0.17
Other Sources <1%	Wikipedia	350	0.02
	Videos	274	0.02
	Q&A posts	3	<0.01
Academic Sources <1%	Peer reviews	299	0.02
	Faculty opinions	38	<0.01
Policy & Patents <1%	Policy documents	165	0.01
	Patents	4	<0.01
Total		1,751,396	100

The highest number of mentions of Covid-19 articles was driven by social media, accounting for about 97% of coverage, while news and blogs accounted for about 3% of mentions. The first recorded article in our dataset was on the 1st of January, and the first mention of Covid-19 articles was recorded on January 14th on Altmetric.com.

Subject network analysis

Nearly two years of pandemic-related article subject network analysis showed that different fields and subjects are involved in addressing pandemic-related challenges. At the start of the pandemic, the majority of the subjects involved were those directly affected by the pandemic (i.e., medical and health sciences). As illustrated in Figure 2, during the first three months of the pandemic, research outputs were heavily concentrated in medical and health sciences, which accounted for more than two-thirds of all nodes. Virology, epidemiology, and clinical medicine were the dominant subjects, strongly interconnected with biological sciences. Psychology and social sciences appeared only at the margins of the network, representing less than 10% of subjects. For example, early studies emphasized rapid detection of the virus and development of diagnostic kits, reflecting the urgent biomedical focus of the global scientific community.

In other words, during the initial months of the pandemic, research publications were predominantly focused on topics such as virology, epidemiology, and clinical sciences. These fields formed the core of early studies as researchers concentrated on understanding the virus,

its transmission, and potential treatment strategies. At this stage, strong connections existed between the biological and health sciences, and the thematic networks were densely concentrated in these areas.

As the pandemic continued and the psychological and social impacts of lockdowns and economic shifts intensified, research topics gradually broadened to include psychology, psychiatry, and social sciences. This period saw a substantial increase in studies related to psychological effects, social pressures, and the impact of isolation. These thematic shifts reflect evolving societal and psychological needs, as well as the scientific community's efforts to address the mental health consequences brought about by the crisis.

In the ensuing months, particularly with the onset of mass vaccination campaigns, themes related to economics, crisis management, and health policy were added to the thematic networks. These topics focused on analyzing and assessing the economic repercussions and management challenges posed by the pandemic, reflecting the scientific community's growing interest in exploring economic and managerial solutions to navigate the crisis more effectively. In the later stages of the pandemic, new topics such as banking, food sciences, and religious studies emerged within the thematic networks. This evolution highlighted the progressively widening scope of health crises, as they began to encompass issues beyond health and into areas not traditionally associated with public health.

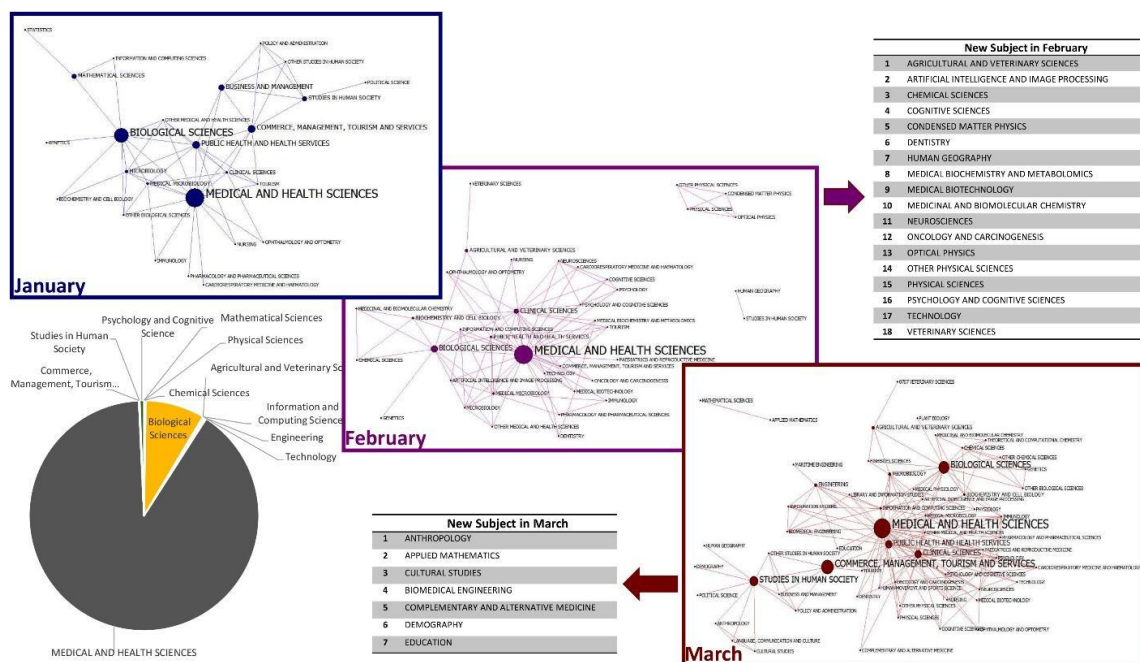


Figure 2: Subject network of the published articles in the first three months of the pandemic. (VOSviewer version 1.6.17 was used to visualize subject networks)

Keywords' co-occurrence networks in January show that published articles are in 25 subject categories. Circles represent the subjects, and links show the relation between subjects. The February keyword co-occurrence network has 38 nodes and shows an increase in the number of subjects (18 new subjects). In this month, five subjects from the January network were eliminated from the February network (i.e., business and management, mathematical sciences, other biological sciences, policy and administration, and statistics). On March 19, new nodes were added to the network, and 3 subjects (medical biochemistry and metabolomics, optical

physics, and condensed matter physics) were removed from it. The total number of nodes in March was 60.

Figure 3 shows the subject network of the published papers in September, October, and November of 2021. By late 2021, the subject network had expanded substantially, with more than 60 nodes covering diverse disciplines. While public health and clinical sciences remained central, new subjects such as economics, banking, food sciences, law, and religious studies became more visible, reflecting the pandemic's multidimensional consequences. For instance, studies in economics examined supply chain disruptions and labor market shocks, while publications in psychology explored the mental health toll of prolonged lockdowns. Legal studies focused on human rights issues during quarantine enforcement, and religious studies addressed the role of faith communities in coping with social isolation. These thematic expansions demonstrate how the pandemic evolved from a primarily biomedical crisis into a broad societal challenge

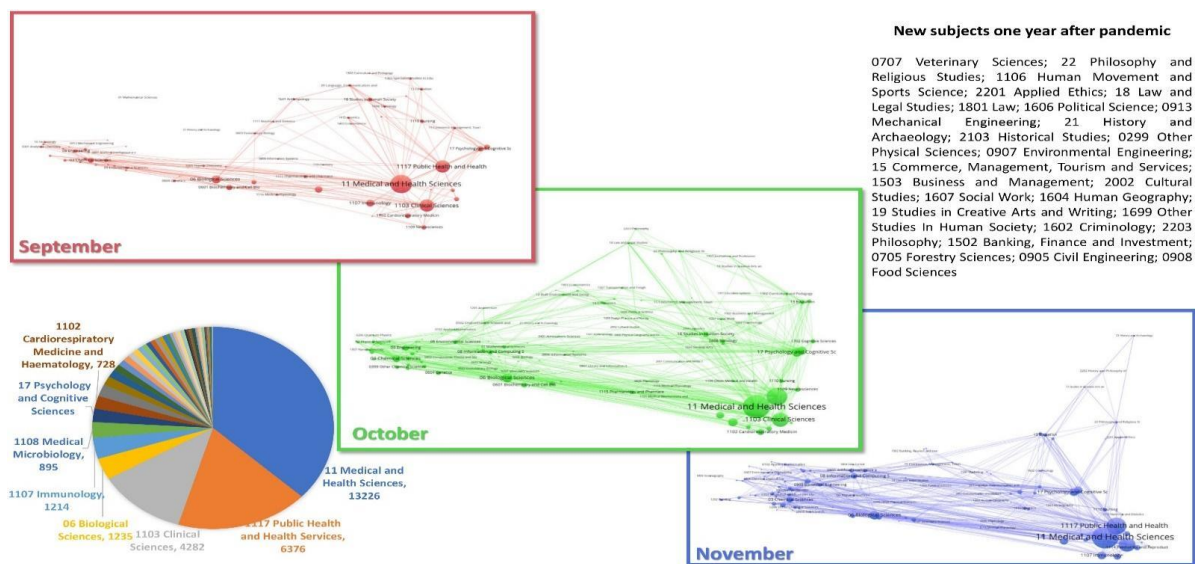


Figure 3: Subject network of the published articles for September, October, and November of 2021. (VOSviewer version 1.6.17 was used to visualize subject networks)

Discussion

The widespread use of social networks during the COVID-19 crisis, combined with societal consequences such as lockdowns, positioned these platforms as critical tools for health communication and even for predicting disease spread and behavioral responses. Our findings reveal that Twitter played the dominant role as the primary source of mentions for COVID-19 research articles, followed by news media as the second most significant channel for disseminating health information. This observation is consistent with comparable studies that have identified Twitter as the main channel for the recirculation of COVID-19 papers, noting that as much as 96% of all mentions for these articles occurred on this platform (Tornberg et al., 2022).

Furthermore, the news media demonstrated notably higher coverage of COVID-19 research than of other health domains (Fang & Costas, 2020). For instance, while only about 11% to 16% of cancer-related articles received media coverage (Haneef, Ravaut, Baron, Ghosn &

Boutron, 2017; Maggio, Ratcliff, Krakow, Moorhead, Enkhbayar & Alperin, 2019), the media coverage for COVID-19 articles in the present study was substantially greater. This aligns with other analyses showing that COVID-19 has been one of the most highly referenced scientific topics in the digital space in terms of social and media attention (Moon et al., 2021).

A key distinguishing feature of this research is its reliance on Altmetric data anchored to verified scientific content (i.e., articles indexed in PubMed), which distinguishes it from many previous studies. A majority of earlier research utilized non-scientific Twitter messages or Google Trends data, sources prone to disseminating misinformation or contributing to infodemics. A systematic review indicated that although 93% of studies based on Google Trends reported a positive correlation between health search behavior and disease monitoring, only 7% provided a reproducible search strategy (Nutti et al., 2014). In contrast, the current study utilized a reproducible data collection strategy, thereby enhancing the validity and credibility of the findings.

The combination of Altmetric data with an analysis of published article content provides valuable insight into the thematic evolution of COVID-19 research. In the initial phases of the pandemic, the focus was overwhelmingly on basic, biological, and clinical sciences. However, following the initiation of mass vaccination campaigns, new topics emerged that spanned economics, transportation, tourism, and even religious studies. This evolution suggests that health crises initially draw the scientific community's attention to biological and medical aspects, but over time, psychological, social, and economic consequences become prioritized in research agendas. A similar pattern has been reported elsewhere, demonstrating that pandemics gradually shift research trajectories from basic sciences toward social sciences and policy studies (Valderrama-Baca et al., 2023).

These findings demonstrate that Altmetrics is not merely a tool for gauging the social impact of scientific articles; it can also reveal temporal and thematic patterns in the evolution of research. Such analysis is invaluable for health policymakers and crisis managers, as identifying past trends can help predict future research needs and social issues, thereby facilitating more appropriate responses to similar crises.

Limitations

Despite the valuable insights, an exclusive reliance on Altmetric data cannot suffice for a comprehensive analysis, as these data primarily reflect public attention and are not necessarily indicative of scientific quality. The high volume of mentions merely reflects visibility, not inherent scientific rigor or impact. Furthermore, our focus was primarily on articles indexed in major databases, potentially overlooking research disseminated through less formalized channels or in non-English languages.

Therefore, the future of research in this domain must move toward integrating quantitative and qualitative approaches, utilizing multilingual and multi-source data, and combining Altmetrics with concrete policy evidence. Only through such integration can the gap between scientific knowledge production and practical decision-making be narrowed, improving preparedness for future crises. The findings of this study remind policymakers that effective management of health crises requires the immediate adoption of an interdisciplinary perspective, multi-level evidence, and social data alongside biological sciences.

Conclusion

The results of this study underscore that social networks and news media during the COVID-19 pandemic transcended their role as mere communication tools, becoming platforms that reflected, amplified, and even steered scientific and public attention. The analysis, which integrates Altmetrics with an examination of the thematic evolution of scholarly articles, demonstrates that public health crises rapidly redefine research priorities: an intense focus on medical and biological sciences in the early stages, followed by an expansion into psychology, economics, policy-making, and even religion and culture in later stages. These findings affirm that pandemics are multidimensional phenomena that affect not only the healthcare system but all facets of social life.

Ethics approval and consent to participate

This study has used metadata and statistical data from authorized sources. This study has been evaluated and found to be in accordance with the ethical principles and the international norms and standards for conducting Medical Research.

Consent for publication

Authors declare consent for publication of the research report.

Availability of data and material

The dataset supporting the conclusions of this article is available in the Zonodo repository

Competing interests

Authors declare no competing interests.

Acknowledgements

The data was collected from Altmeteric.com by access provided through Simon Fraser University's Scholarly Communication Lab.

References

- Ahmed, R. & Mao, Y. (2022). Communication research on health disparities and coping strategies in COVID-19-related crises. *Health Communication*, 37(12), 1455-1456. <https://doi.org/10.1080/10410236.2022.2111633>
- Bundy, J., Pfarrer, M., Short, C. & Coombs, W. (2017). Crises and crisis management: integration, interpretation, and research development. *Journal of Management*, 43(6), 1661-1692. <https://doi.org/10.1177/0149206316680030>
- Elmore, S. A. (2018). The altmetric attention score: what does it mean and why should I care? *Toxicologic Pathology*, 46(3), 252-255. <https://doi.org/10.1177/0192623318758294>
- Fang, Z. & Costas, R. (2020). Tracking the Twitter attention around the research efforts on the COVID-19 pandemic. <https://doi.org/10.48550/arXiv.2006.05783>
- Furstenau, L. B., Rabaioli, B., Sott, M. K., Cossul, D., Bender, M. S., Farina, E. M. J. D. M., ... & Bragazzi, N. L. (2021). A bibliometric network analysis of coronavirus during the first eight months of COVID-19 in 2020. *International Journal of Environmental Research and Public Health*, 18(3), 952. <https://doi.org/10.3390/ijerph18030952>

- Haghani, M., Bliemer, M. C., Goerlandt, F. & Li, J. (2020). The scientific literature on Coronaviruses, COVID-19, and its associated safety-related research dimensions: A scientometric analysis and scoping review. *Safety Science*, 129, 104806. <https://doi.org/10.1016/j.ssci.2020.104806>
- Haneef, R., Ravaud, P., Baron, G., Ghosn, L. & Boutron, I. (2017). Factors associated with online media attention to research: A cohort study of articles evaluating cancer treatments. *Research Integrity and Peer Review*, 2, 9. <https://doi.org/10.1186/s41073-017-0033-z>
- Ioannidis, J. P., Cripps, S. & Tanner, M. A. (2022). Forecasting for COVID-19 has failed. *International Journal of Forecasting*, 38(2), 423-438. <https://doi.org/10.1016/j.ijforecast.2020.08.004>
- Maggio, L. A., Ratcliff, C. L., Krakow, M., Moorhead, L. L., Enkhbayar, A. & Alperin, J. P. (2019). Making headlines: an analysis of US government-funded cancer research mentioned in online media. *BMJ open*, 9(2), e025783. <https://doi.org/10.1136/bmjopen-2018-025783>
- Mardiani, E., Riniati, W. O., Mokodenseho, S., Podomi, A. & Mamonto, J. (2024). Bibliometric analysis of hybrid and distance learning literature in the digital age. *The Eastasouth Journal of Learning and Education*, 2(01), 12-20. <https://doi.org/10.58812/esle.v2i01.229>
- Moon, J. Y., Yoon, D. Y., Hong, J. H., Lim, K. J., Baek, S., Seo, Y. L. & Yun, E. J. (2021). The most widely disseminated COVID-19-related scientific publications in online media: A bibliometric analysis of the top 100 articles with the highest altmetric attention scores. *Healthcare (Basel, Switzerland)*, 9(2), 239. <https://doi.org/10.3390/healthcare9020239>
- Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I. & Murugiah, K. (2014). The use of Google Trends in health care research: A systematic review. *PloS ONE*, 9(10), e109583. <https://doi.org/10.1371/journal.pone.0109583>
- Ozuem, W., Willis, M. & Howell, K. (2022). Thematic analysis without paradox: Sensemaking and context. *Qualitative Market Research: An International Journal*, 25(1), 143-157. <https://doi.org/10.1108/QMR-07-2021-0092>
- Riley, C., Xie, B. & Khurshid, A. (2021). Challenges encountered in comparing international policy responses to COVID-19 and their effects. *Health Research Policy and Systems*, 19(1), 134. <https://doi.org/10.1186/s12961-021-00783-1>
- Tornberg, H., Moezinia, C., Wei, C., Bernstein, S. A., Wei, C., Al-Beyati, R., Quan & Diemert, D. J. (2023). Assessment of the dissemination of COVID-19-related articles across social media: Altmetrics study. *JMIR Formative Research*, 7(1), e41388. <https://doi.org/10.2196/41388>
- Valderrama-Baca, P., Arroyo-Machado, W. & Torres-Salinas, D. (2023). Altmetrics can capture research evidence: A study across study types in the COVID-19 literature. <https://doi.org/10.48550/arXiv.2303.09697>
- Vosoughi, S., Roy, D. & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>
- Zarocostas, J. (2020). How to fight an infodemic. *The Lancet*, 395(10225), 676. [https://doi.org/10.1016/s0140-6736\(20\)30461-x](https://doi.org/10.1016/s0140-6736(20)30461-x)