

Review

Public Health Using Social Network Analysis During the COVID-19 Era: A Systematic Review

Stanislava Gardasevic ¹, Aditi Jaiswal ^{2,*}, Manika Lamba ³, Jena Funakoshi ⁴, Kar-Hai Chu ⁵, Aekta Shah ⁶, Yinan Sun ¹, Pallav Pokhrel ⁷ and Peter Washington ^{2,*}

¹ Communication and Information Sciences, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA

² Information and Computer Sciences, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA

³ School of Library & Information Studies, University of Oklahoma, Norman, OK 73019, USA; manika@ou.edu

⁴ Office of Public Health Studies, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA

⁵ Behavioral and Community Health Sciences, Clinical and Translational Science Institute, University of Pittsburgh, Pittsburgh, PA 15260, USA

⁶ Unaffiliated Researcher, Oakland, CA 94612, USA

⁷ University of Hawai'i Cancer Center, University of Hawai'i at Mānoa, Honolulu, HI 96822, USA

* Correspondence: ajaiswal@hawaii.edu (A.J.); pyw@hawaii.edu (P.W.)

Abstract: Social network analysis (SNA), or the application of network analysis techniques to social media data, is an increasingly prominent approach used in computational public health research. We conducted a systematic review to investigate trends around SNA applied to social media data for public health and epidemiology while outlining existing ethical practices. Following PRISMA guidelines, we reviewed articles from Web of Science and PubMed published between January 2019 and February 2024, leading to a total of 51 papers surveyed. The majority of analyzed research (69%) involved studying Twitter/X, followed by Sina Weibo (16%). The most prominent topics in this timeframe were related to COVID-19, while other papers explored public health topics such as citizen science, public emergencies, behavior change, and various medical conditions. We surveyed the methodological approaches and network characteristics commonly employed in public health SNA studies, finding that most studies applied only basic network metrics and algorithms such as layout, community detection, and standard centrality measures. We highlight the ethical concerns related to the use of social media data, such as privacy and consent, underscoring the potential of integrating ethical SNA with more inclusive, human-centered practices to enhance the effectiveness and community buy-in of emerging computational public health efforts.

Keywords: public health; epidemiology; COVID-19; social network analysis; social media



Citation: Gardasevic, S.; Jaiswal, A.; Lamba, M.; Funakoshi, J.; Chu, K.-H.; Shah, A.; Sun, Y.; Pokhrel, P.; Washington, P. Public Health Using Social Network Analysis During the COVID-19 Era: A Systematic Review. *Information* **2024**, *15*, 690. <https://doi.org/10.3390/info15110690>

Academic Editor: Neil Vaughan

Received: 24 September 2024

Revised: 25 October 2024

Accepted: 26 October 2024

Published: 2 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Social media platforms allow users to share information, socialize, and collaborate [1]. In 2023, an estimated 4.9 billion people used social media globally, with the average user maintaining 8.4 social accounts and spending approximately 145 minutes per day on these platforms [2]. The increasing popularity of social media has rendered these platforms as useful tools for public health researchers, offering a powerful means both to analyze network effects related to health surveillance and to disseminate health-related information and education among target populations [3] and the general public. However, the pervasiveness of social media has also facilitated the spread of misinformation on several health issues, including vaccines, substance use, and diseases [4], negatively impacting individuals and communities [5]. During health crises such as the COVID-19 pandemic, misinformation about the virus, preventive measures, and treatments were widely circulated, causing confusion and harmful behaviors [6].

Social network analysis (SNA) is a popular method of studying such phenomena on social media. SNA consists of the application of network analysis algorithms and concepts

to social media data. SNA involves the investigation of the connections and relationships between actors within a social network to understand the network's structures and features, interaction patterns, and influential nodes [7,8]. This approach typically involves visualizing connections (edges) between actors (nodes) in a network graph followed by applying mathematical and statistical methods for further analysis. Historically, SNA has been employed in epidemiological research to trace the spread of diseases by mapping the interactions and connections among individuals and communities. This allows researchers to identify key influencers, understand information dissemination patterns, monitor real-time trends in disease-related behaviors, and enhance surveillance strategies for early outbreak detection [9,10].

Because social media platforms rely on a foundation of user relationships, SNA has emerged as a natural methodology for analyzing social media datasets. SNA can be used to investigate health-related information dissemination and identify influential entities within public forums, offering insights into connections and interactions among individuals and communities. In such research, the networks are typically created as directed networks based on platform user interactions (e.g., mentions, retweets, likes) or as undirected thematic networks based on topics discussed (e.g., hashtags as nodes, linked if they appear in the same tweet). Table 1 provides concise definitions of primary SNA approaches and introduces relevant algorithms and metrics for the study.

Table 1. Overview of common social network analysis methods.

Family of Methods	Description	Examples
Layout algorithms	These algorithms aim to produce a visually useful arrangement of nodes and edges in a graphical display. This arrangement is designed to visually depict relationships in a way that is easy to understand and analyze, highlighting patterns like clusters, central figures, and isolated nodes [11].	Force-directed Layouts are dynamic layouts where nodes repel each other like charged particles, while edges act like springs holding nodes together. This results in a layout that naturally separates clusters and displays tightly knit groups closer together. For example, ForceAtlas2 is specifically designed for the visualization of large graphs, enhancing readability by distributing nodes in a manner that reflects their relational metrics, emphasizing clustering based on node interactions [12].
Community detection algorithms	These algorithms cluster nodes within a broader network, guided by the relationships or interactions among the nodes. This approach seeks to uncover the inherent social structures within the network, thereby enhancing our comprehension of information dissemination and the formation of influential groups within the network [13].	Modularity Optimization is used to detect community structure in networks by optimizing modularity, a metric that quantifies the strength of the division of a network into communities. For example, the Louvain Method iteratively optimizes community structure for maximum modularity [14].
Centrality metrics	These metrics aim to identify the most important nodes within a graph. Centrality metrics are crucial for understanding the influence of particular nodes in network structures [15].	Betweenness Centrality measures the extent to which a node lies on paths between other nodes. It quantifies the number of times a node acts as a bridge along the shortest path between two other nodes [14].

SNA has the potential to improve public health communication strategies by providing insights beyond simple analytics metrics such as the number of retweets, followers, or likes on Twitter/X [16]. When combined with techniques like topic modeling and sentiment analysis, SNA enables researchers to monitor real-time trends in public health behaviors and sentiments [17].

Despite its widespread use in public health research, significant gaps remain in the ethical frameworks governing social media-based research, lacking universally agreed-upon norms or best practices [18]. The inability of users to easily delete their data once they are scraped off social media and in the hands of researchers poses a risk to user privacy [19], consent, and data ownership [20]. In most studies, the user data are typically safeguarded through anonymizing social media data by removing personally identifiable information (PII) such as user IDs, names, age, and location information [21]. Obtaining informed consent from the users [19] is often deemed impractical considering the volume of data often collected [22] and the methodological limitations that the reduction in sample size can lead to [23]. With the ongoing academic debate surrounding the ethics and bias of social media analysis research, especially in light of the Cambridge Analytica scandal [24], there has been a significant reduction in the availability of social media data to researchers. API access to user data has been restricted on many platforms to prevent any potential misuse of user data. This puts more responsibility on the researchers to ethically use the data while safeguarding participant information, particularly in health-related contexts [25]. Another critical issue is the lack of representativeness in social media data. It is important to recognize that social media usage differs by socioeconomic demographics and the user base may not accurately represent the broader population [26]. This has the potential to introduce sampling biases, constraining the applicability of research outcomes or harming marginalized populations.

Given these considerations, we sought to systematically review the state of scientific inquiry utilizing SNA on social media data in the domain of public health and epidemiology using PRISMA guidelines, with a critical eye towards ethical best practices. We sought answer to the following research questions (RQs):

- (RQ1) What social media platforms are used to collect data relevant to public health and epidemiology?
- (RQ2) What are the health domains that these analytical techniques are commonly applied to?
- (RQ3) What are the prevalent structural characteristics of the networks created from social media platform data, and what SNA methods and software are applied to study these networks?
- (RQ4) What are the ethical trends and concerns around using social media data for SNA research for public health?

While previous reviews have examined the use of SNA in broader health contexts, this study focuses on the application of SNA methodologies during the COVID-19 era, leading to an enrichment of COVID-related content in this review. Thus, we highlight modern public health challenges such as misinformation and vaccine debates that were not previously central to healthcare-related SNA reviews.

2. Methods

2.1. Search Strategy and Selection of Studies

The data collection process (Figure 1) for this systematic literature review adhered to the PRISMA guidelines. We searched the Web of Science and PubMed databases in March 2024. We performed the initial informal search, which was followed by iterative keyword searches within both databases to identify relevant papers. The search string encompassed terms related to “Social Media”, “Social Networking”, “Social Network”, “Social Networks”, “Social Network Analysis”, “Network Science”, “Public Health”, and “Epidemiology”, due to their direct relevance to the chosen topic of study.

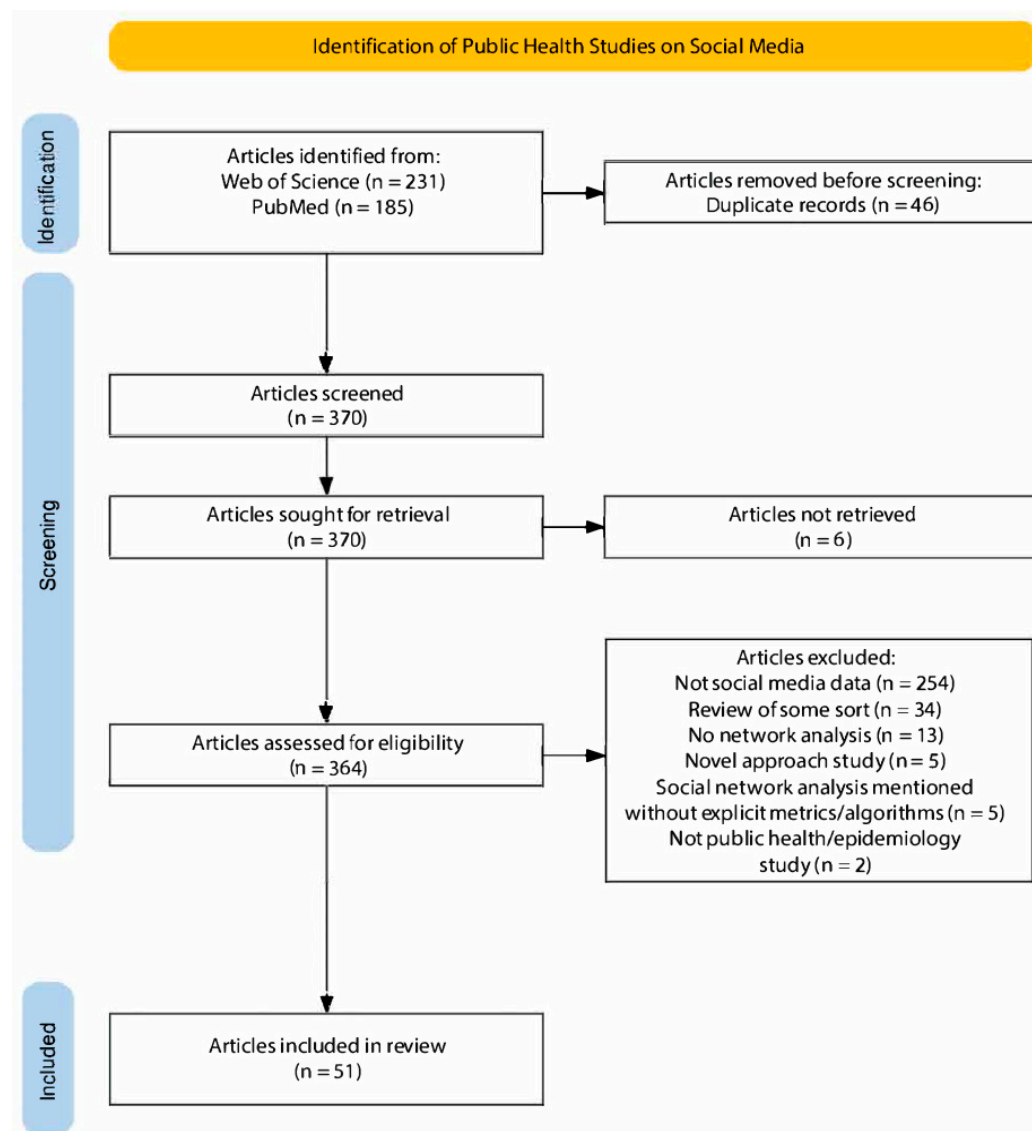


Figure 1. Identification of public health literature related to SNA using PRISMA guidelines.

Specifically, the search query we used for the Web of Science database was: (TS = “Social Media” OR TS = “Social networking” OR TS = “Social network” OR TS = “Social networks”) AND (TS = “Social network analysis” OR TS = “Network science”) AND (TS = “Public Health” OR TS = “Epidemiology”). We restricted the query to include journal articles published in English between 1 January 2019 and 29 February 2024.

Similarly, in PubMed, our search strategy involved Medical Subject Headings (MeSH) terms and text words. The detailed search string was: ((“Social Media”[MeSH Terms] OR “Social Media”[Text Word] OR “Social networking”[MeSH Terms] OR (“Social network”[Text Word] OR “Social networks”[Text Word] OR “Social networking”[Text Word])) AND (“social network analysis”[Text Word] OR “network science”[Text Word]) AND (“public health informatics”[MeSH Terms] OR “epidemiology”[MeSH Subheading]) AND (“public health”[MeSH Terms] OR (“public”[All Fields] AND “health”[All Fields]) OR “public health”[All Fields])). We restricted the query to include only full-text journal articles published in English between 1 January 2019 and 29 February 2024.

We utilized the PICO Portal to analyze articles based on predefined inclusion and exclusion criteria. Initially, we screened titles and abstracts screened for eligibility. If a paper’s abstract was considered irrelevant, it was excluded from further analysis. Conversely, if abstracts were deemed relevant, the full text of the paper underwent eligibility review. We

discarded the papers meeting exclusion criteria. Reviewers were divided into groups, with each paper reviewed by two reviewers. Papers with discrepancies or coded for inclusion as “maybe” were revisited by four reviewers to reach a consensus.

2.1.1. Inclusion Criteria

We set stringent inclusion criteria for the systematic review. Studies had to be reported in English and published within the last five years, between 1 January 2019 and 29 February 2024. Additionally, all studies needed to focus on the utilization of social media data subjected to SNA and be peer-reviewed scholarly journal articles without demographic restrictions. The scope was confined to the fields of public health and epidemiology.

2.1.2. Exclusion Criteria

We excluded the studies that did not employ SNA, did not utilize social media platforms’ data, or were not relevant to the field of public health or epidemiology. Additionally, we excluded studies lacking explicit mentions of the metrics and algorithms used in the SNA of social media data. Furthermore, we excluded reviews or meta-analyses, such as scoping reviews, systematic reviews, or bibliometric/citation analyses.

3. Results

A total of 416 records were initially identified through the search process, of which we removed all the duplicate records ($n = 46$), leaving us with 370 unique records for consideration based on their titles and abstracts. We excluded articles if their titles or abstracts did not contain one or more of the search terms utilized. Additionally, papers for which we could not access the full text were excluded at this stage ($n = 6$). Subsequently, we evaluated abstracts for eligibility and excluded papers if they were (i) reviews, (ii) did not utilize social media data or SNA, or (iii) were not within the field of public health/epidemiology, broadly defined. The remaining articles, deemed eligible, underwent full-text review, resulting in a total of 51 studies eligible for inclusion, as illustrated in Figure 1.

Tables 2–4 list the papers we included in the systematic review, outlining the categories pertinent to our research questions; detailing the different topics discussed on each platform; showing information on network directionality, layout, community detection, centrality algorithms, and other metrics used for analysis; and identifying the ethical considerations relevant to each publication. We grouped the papers into tables based on the (i) social media platform and (ii) public health topic examined.

Table 2. Research papers using Twitter/X data.

Public Health Topic: COVID-19 Pandemic	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
Social response and key actors (US, Canada, Mexico)	Yum, 2023 [27]	Directed	N/A	Clauset–Newman–Moore	In-degree, betweenness, eigenvector	N/A	N/A
Sentiments (topics)	Hung et al., 2020 [28]	Undirected	N/A	N/A	Degree, closeness, betweenness, eigenvector	N/A	N/A
Evolutionary trend of information spread	Xing et al., 2021 [29]	Directed	ForceAtlas 2; Yifan Hu Multilevel	Clustering coefficient	Degree, weighted degree, in-degree, out-degree eigenvector	Average path length, k-core analysis	N/A
Pandemic and food poverty topics	Eskandari et al., 2022 [30]	Undirected	Network visualization (unspecified)	Unspecified	N/A	N/A	Anonymized users
Discourse on science and health	Durazzi et al., 2021 [31]	Directed	ForceAtlas 2	Louvain	Out-degree	Largest connected component	N/A
Discourse & influencers	Yum, 2021 [32]	Directed *	N/A	Clauset–Newman–Moore	Betweenness	N/A	N/A
Masks discourse-influential users	Ahmed et al., 2020 [33]	Directed *	Harel–Koren Fast Multiscale	NodeXL clustering based on mentions	Betweenness	N/A	Anonymized users (that are not public figures/organizations)
Public discourse	Pascual-Ferrá et al., 2020 [34]	Directed *	Distributive Recursive Layout (DrL)	Modularity	Degree in-degree, out-degree	Diameter, density, reciprocity,	N/A
Public discourse (Italy)	Massaro et al., 2021 [35]	Directed	ForceAtlas 2	Unspecified	N/A	N/A	N/A
Health condition: COVID-19 misinformation	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
Plandemic misinformation campaign	Nazar and Pieters, 2021 [36]	Directed	OpenOrd, ForceAtlas 2, Yifan Hu Multilevel	Modularity	In-degree, out-degree, betweenness	N/A	Anonymized users, redacted links
Bots vs. human, communication patterns	Weng and Lin, 2022 [37]	Directed	Gephi visualization (unspecified)	N/A	In-degree, out-degree, betweenness	N/A	N/A

Table 2. Cont.

Public Health Topic: COVID-19 Pandemic	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
COVID-19 and 5g conspiracy	Ahmed et al., 2020 [38]	Directed	Harel-Koren Fast Multiscale	Clauset–Newman– Moore	Betweenness	N/A	Anonymized users
Conspiracy theory	Gruzd and Mai, 2020 [39]	Directed *	Gephi visualization (unspecified)	Fast Unfolding community detection	N/A	N/A	N/A
Information diffusion (including miss/disinformation)	Kothari et al., 2022 [40]	Directed *	N/A	Modularity	Degree	Diameter, degree of distinction among network clusters	Anonymized users
Misinformation in Pennsylvania	Alieva et al., 2023 [41]	Directed *	N/A	Louvain	Degree	N/A	N/A
Spread of misinformation	Duzen et al., 2023 [42]	Directed *	N/A	Girvan-Newman	Degree, in-degree, out-degree, closeness, betweenness	N/A	Anonymized users
Conspiracy theory & bots	Ahmed et al., 2020 [43]	Directed	Harel-Koren Fast Multiscale	Clauset–Newman– Moore	In-degree, out-degree, betweenness	N/A	N/A
Health condition: COVID-19 Vaccines	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
Vaccine Debate in Turkiye	Seckin et al., 2024 [44]	Directed	N/A	Louvain	In-degree	N/A	N/A
Vaccine misinformation in Turkiye	Durmaz and Hengirmen, 2022 [45]	Directed	N/A	N/A	In-degree, betweenness	N/A	Changed names of users
Vaccine trial- opinions and behaviors	Boucher et al., 2021 [46]	Directed	N/A	Louvain	N/A	N/A	Anonymized users
Vaccination in Chile	Villegas et al., 2022 [47]	Directed	N/A	Clustering coefficient	Betweenness, PageRank	N/A	N/A
Misinformation on AZ vaccine & blood clot	Hobbs et al., 2024 [48]	Direcred, Undirected	ForceAtlas 2	N/A	In-degree, out-degree, eigenvector	N/A	Anonymised datasets
Magnetic Vaccine Conspiracy Theory	Ahmed et al., 2023 [49]	Directed	ForceAtlas 2	Clauset–Newman– Moore	Betweenness	N/A	Top 4 account not anonymized for accountability
Health condition: Other	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
Brain tumor	Feliciano et al., 2020 [50]	Directed *	Gephi visualization (unspecified)	N/A	Out-degree	N/A	N/A

Table 2. Cont.

Public Health Topic: COVID-19 Pandemic	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
79 conditions	Zhang and Ahmed, 2019 [51]	Directed	Harel–Koren fast multiscale	Clauset–Newman–Moore	N/A	N/A	N/A
Breastfeeding	Moukarzel et al., 2021 [16]	Directed *	Modularity *	Modularity *	Degree	N/A	Dataset is not public due to the threat of deanonymization by copy/pasting the text
Ebola	Liang et al., 2019 [52]	Directed *	N/A	N/A	Degree, authority	N/A	N/A
Turkiye earthquake response	Aktuna and Bahar-Özvarış, 2023 [53]	Directed *	N/A	Clauset–Newman–Moore	In-degree, out-degree, eigenvector, betweenness	N/A	N/A
HPV vaccine discussions in COVID-19 times	Boucher et al., 2023 [54]	Directed *	N/A	Louvain	N/A	N/A	N/A
World Breastfeeding Week	Moukarzel et al., 2021 [55]	Directed *	Modularity	Modularity	Degree, in-degree, out-degree, betweenness,	N/A	Data is available for other researchers only upon request to protect the users' privacy
Stroke Awareness Month	Goyal et al., 2022 [56]	Unirected *	ForceAtlas2	N/A	N/A	N/A	N/A
Asthma	Pratiwi et al., 2024 [57]	Directed	Fruchterman-Reingold	N/A	In-degree, out-degree, betweenness	N/A	Anonymized users
Citizen science in health	Martínez-Martínez et al., 2023 [58]	Directed	k-core graph decomposition	Louvain	In-degree, out-degree, Betweenness	N/A	Non-project, institution or official organization username anonymized
Melanoma	Gomaa et al., 2022 [59]	Directed *	Harel–Koren fast multiscale	Clauset–Newman–Moore	In-degree, out-degree, betweenness	Edge weight-reciprocity between users	N/A
Obesity-Related Behavior Change	Park, 2022 [60]	Directed	OpenOrd, Fruchterman Reingold	N/A	N/A	Edge weight-number of interactions	Anonymized users, modified quotations

* This information was not explicitly stated in the paper but was instead implied.

Table 3. Research papers using Sina Weibo data.

Health Condition COVID-19 Misinformation	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
Bots vs. human, communication patterns during COVID-19 Rumors during COVID-19	Cai et al., 2022 [61]	Directed	UMAP optimized layout	N/A	Degree, Intermediate PageRank	N/A	N/A
	Lv et al., 2022 [62]	Undirected	N/A	N/A	Degree	N/A	N/A
Health condition: COVID-19 Vaccines	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
COVID-19 Vaccine Information Forwarding	Sun et al., 2022 [63]	Directed	N/A	N/A	Degree, closeness, betweenness	Average distance, diameter, cohesion	N/A
COVID-19 Vaccines	Wang et al., 2022 [64]	Directed	Fruchterman Reingold	Degree of homophily	In-degree, out-degree	N/A	Anonymized users
Health condition: Other	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
Topic and emotion propagation in public emergencies	Cai et al., 2021 [65]	Directed	Fruchterman Reingold	Modularity	Degree, PageRank, network diameter, eigenvector	Diameter, Average path length	N/A
Depression	Shi and Khoo, 2023 [66]	Directed	N/A	Louvain	Degree in-degree, out-degree, betweenness	Global efficiency, average clustering coefficient	N/A
Public sentiment on emergencies	Li et al., 2020 [67]	Directed	N/A	N/A	Degree	Density	N/A

Table 4. Research papers using multiple or other platforms (health conditions ordered alphabetically).

Platform(s)	Health Condition	Paper	Directed/ Undirected	Layout	Community Detection	Centrality	Other Metrics	Ethical Implications
Facebook	Anti-vaccination sentiments	Hoffman et al., 2019 [68]	Undirected *	Gephi visualization (unspecified)	Modularity	Degree, betweenness	Density	N/A
Twitter; Sina Weibo	COVID-19 public concern discourse	Deng and Yang, 2021 [69]	Directed *	UCINET 6 visualization (unspecified)	N/A	Degree, betweenness	N/A	N/A
Twitter, Reddit, YouTube	COVID-19 vaccine side effects	Kwon and Park, 2023 [70]	Undirected	Fruchterman Reingold	Louvain	N/A	Edges weight-Jaccard similarity score	Anonymized data, Modified quotations
YouTube	COVID-19 Anti-Vaccine Videos	Gruzd et al., 2023 [71]	Undirected	N/A	Fast Unfolding community detection, Newman's modularity class detection algorithm	In-degree	LOLOG (Latent Order Logistic)	N/A
WeChat	Health (injury) education for parents of children 0–3 yo (China)	Feng et al., 2023 [72]	Directed	N/A	N/A	Centrality, point centrality, betweenness, closeness	Density, distance, coreness, basic ego measure, cohesion subgroup (components, cliques, structural hole)	N/A
Baidu Tieba	Online psychological and physiological disease communities	Yao et al., 2022 [73]	Directed *	N/A	Louvain, average clustering	Degree, in-degree, out-degree, closeness, betweenness	Coefficient, average path length, weight,	N/A
Baidu Tieba	Public and government participation during COVID-19 Smoking, Hypertension, Vaccination, Nutrition influence maximization	Fu et al., 2022 [74]	Undirected *	Agglomeration subgroups	N/A	Centrality	Density	N/A
Twitter, Sina Weibo		Ahmed and Lugovic, 2019 [75]	Directed	N/A	N/A	Degree, authority	Cascade size, scale and depth	N/A

* This information was not explicitly stated in the paper but was instead implied.

3.1. (RQ1) Social Media Platforms

We found that the majority of papers reviewed (65%; 38/51) used Twitter/X as the primary platform of study. The next most common platform was Sina Weibo (20%; 10/51), followed by Baidu Tieba (4%; 2/51), Facebook (2%, 1/51), WeChat (2%, 1/51), Reddit, and YouTube (2%, 1/51), all being used separately or in combination with each other [46,70,71].

3.2. (RQ2) Health Domains

As shown in the Tables 2–4, the majority (63%; 32/51) of the reviewed papers in the period from 2019 to mid 2024 were dedicated to topics related to COVID-19 such as (i) public communication and perception of COVID-19 (22%; 11/51), (ii) mis/disinformation spread and conspiracy theories (16%; 8/51 papers), (iii) vaccine-related topics (20%; 10/51 papers), identifying bots vs. humans posting about vaccines (6%; 3/51 papers), and (iv) the intersection of COVID-19 and food poverty (2%; 1/51). This distribution can be attributed to the timeline of this review: the pandemic and post-pandemic era.

A few other papers focused on social media discussions and sentiments around non-COVID-19 topics such as public emergencies [65,67], including earthquake responses [53]; HPV vaccines [54]; breastfeeding [16,55]; anti-vaccination sentiments [68]; health conditions, such as brain tumor [50], Ebola [52], stroke [56], asthma [57], melanoma [59], psychological and physiological disease [73], and depression [66]; social media education for parents of children (0–3 years) around injuries [72]; obesity-related behavior change [60]; citizen science for health care [58]; and smoking, hypertension, vaccination, and nutrition influence maximization [75]. These results collectively illustrate how SNA can be applied to a broad range of health conditions and public health inquiries.

3.3. (RQ3) Structural Characteristics of Networks and Analytical Methods

The majority of the reviewed studies performed SNA on directed networks (84%; 43/51). While some explicitly stated this, for others (annotated with * in Tables 2–4), the structural characteristic of the network was inferred based on the narrative around network analytics.

Networks created from Twitter/X data typically operate as directed networks, meaning that in visualizations of Twitter/X data, connections (links or edges) between users (nodes) are often created when one user interacts with another (retweets, mentions, replies), rather than when they mutually follow each other. This differs from Facebook, which functions as an undirected network, where connections between “friends” are represented as ties or edges [34]. Among the reviewed studies, those that performed analytics on the undirected networks were predominantly thematic analyses, where topics or hashtags (nodes) discussed in social media posts were linked based on their joint occurrences.

The reviewed Twitter/X studies tended to capture a subsection of the entire platform, usually harvested through a set of topic-pertinent hashtags in a given time frame. By analyzing these networks to uncover their structural characteristics based on user interactions, researchers can identify key influencers within the community, understand the outreach of the discourse, and examine how users cluster around specific topics or interests. These methods illustrate the broader social dynamics at play and the potential impact of information dissemination on public health behaviors.

Most of the studies applied at least one type of content analysis (qualitative, quantitative, or machine learning) in addition to often different types of network analytics and visualization algorithms to complement their findings. The most commonly used analyses included sentiment, topic, temporal, and regression analyses.

To visually analyze the extracted networks, many studies utilized NodeXL Pro [76], a SNA tool developed by the Social Media Research Foundation. Although the tool previously included the free Twitter search API, this feature has since been discontinued. Other tools utilized to visualize and analyze networks across the corpus included Gephi, UCINET 6, VOSViewer, NetDraw, Netlytic, SocioViz, UCINET, and Python packages such

as NetworkX, Networkit, and Twarc, while the last one was used to check the dataset for deleted and suspended accounts.

The spread of COVID-19 misinformation on social media led to the emergence of the terms ‘infodemic’ and ‘infodemiology’ [46]. Given the suspicion that bots contribute to these discussions, some studies employed Botometer, a machine-learning platform, to compute a bot likeliness ranking for Twitter accounts.

The majority of the papers covered by this review have applied network visualization as a strategy to inspect the community structures that are produced in online discourse in order to highlight key actors and relationships within the network, utilizing a combination of centrality metrics to emphasize prominent users and interactions along with community detection algorithms to distinguish different groups within the network.

Layout algorithms are used to visually represent nodes and edges, making community structures and relationships clear for social network data in public health studies. ForceAtlas 2 and Fruchterman Reingold were the most commonly used layout algorithms, applied in 14% (7/51) and 12% (6/51) of studies, respectively, but not exclusively. The Harel–Koren Fast Multiscale layout algorithm was also frequently used (12%, 6/51) for converting large, complex networks into more understandable visual forms. Other algorithms utilized included Yifan Hu Multilevel [29,36], OpenOrd [34,60], and ForceAtlas [77].

Community detection is used for identifying clusters of users or entities that interact more frequently with each other than with those outside their group. The Louvain algorithm was the most frequently used method (18%, 9/51) and is favored for its efficiency in optimizing modularity to identify dense clusters of interactions such as user groups on social media discussing similar health topics. The second most common community detection method was the Clauset–Newman–Moore algorithm (16%, 8/51) for measuring the fragmentation of a network into distinct communities, where a higher value indicates greater separation between communities [34]. Less frequently used methods like Fast Unfolding (4%, 2/51) and Newman’s modularity class detection (2%, 1/51) cater to large-scale and specialized analyses, respectively, emphasizing computational efficiency and highlighting the diversity in methodological approaches depending on the network size and the specific research questions.

Layout and community detection algorithms selected for the network visualization analytics are often attributed to the choice of software chosen for the analysis (e.g., studies that used NodeXL often applied the readily available Clauset–Newman–Moore clustering algorithm [27,32,38,43,49,51,53,59] and the Harel–Koren Fast Multiscale layout algorithm [33,51,59]).

Centrality metrics are used for identifying influential nodes in social networks. Betweenness centrality was the most utilized metric (45%, 23/51), emphasizing its role in identifying nodes that serve as bridges within the network. In-degree (39%, 20/51), general degree (35%, 18/51) and out-degree centrality (33%, 17/51) were also commonly used to measure the influence based on the number of connections. Other important metrics included eigenvector centrality (13%, 7/51), closeness centrality (10%, 5/51), and PageRank (6%, 3/51). Among other metrics that were reported in the studies, the most prominent were density (12%, 6/51), average path (8%, 4/51), edge weight (8%, 4/51), and diameter (8%, 4/51).

3.4. (RQ4) Ethical Use of Social Media Data

We reviewed the selected papers for potential ethical considerations, focusing on mentions of keywords such as ‘privacy’, ‘ethic*’, and ‘anonym*’. References to Institutional Review Board (IRB) statements were excluded, as some studies had protocol approval while others were exempt due to the public nature of the social media data analyzed.

While ethical considerations are increasingly recognized for research involving social media data, only 17 (33%) of the reviewed studies explicitly addressed ethical concerns. These studies highlighted several key ethical practices, such as the anonymization/de-identification of usernames and other personally identifiable information (PII) to protect

user privacy. Some researchers took meticulous steps to anonymize data, ensuring that users' identities remained confidential, especially those not affiliated with public institutions [30,33,36,38,40,42,46,48,57,58,60], while others emphasized the additional measurement of protection for the social network users by altering direct quotes from social media posts to prevent reverse identification via web searches [60,70]. Still, some studies stated that their anonymization efforts were applied only to individuals, meaning not directed towards projects, institutions, organizations [33,58], or public figures [30,33]. Such measures are crucial in maintaining the balance between leveraging rich social media data and upholding ethical research standards.

The ethical landscape of health-related SNA also involves the handling of misinformation and the accountability of users spreading harmful content. Notably, one study did not anonymize the top four Twitter/X accounts spreading conspiracy theories in order to hold them accountable for their actions [49]. This decision highlights a critical ethical dilemma in public health research: balancing user privacy with public accountability. The ethical considerations extend beyond anonymization and privacy concerns to include issues of informed consent and data ownership, which were sparsely discussed in the reviewed papers. While obtaining informed consent from users on social media may become impractical due to the sheer volume of data, there still lies some ethical responsibilities on the researchers to adhere to platform terms of service and legal frameworks governing data use at a very minimum. These findings collectively highlight the evolving ethical considerations in the field and the need for standardized guidelines in utilizing social media for SNA in public health.

4. Discussion

4.1. Social Media Platforms

Until recently, Twitter provided free access to its data through a user-friendly API, offering a wealth of metadata such as user information, which researchers leveraged to study numerous public health topics. The platform's diverse user-interaction features, such as retweets, follows, quotes, likes, and shared hashtags, enabled several types of SNA studies. However, following its rebranding to X, data access has been restricted and is now primarily available behind a paywall. This change is likely to constrain research to teams with sufficient resources to purchase the data, potentially limiting the scope and inclusivity of future studies.

Although relying on a single social media platform provides an incomplete picture of public opinion, it is important to recognize that different platforms occupy different niches when it comes to information exchange [70]. Each platform also caters to audiences with different demographic backgrounds [78], limiting the ability to generalize findings across the broader population. Unfortunately, only a few studies have conducted comparative analyses across multiple platforms, highlighting a gap in the existing literature [69,70,75].

4.2. COVID-19 Analytics

While COVID-19-related subtopics dominated the literature during the reviewed timeframe, the methodologies and insights drawn from it extend beyond the pandemic, offering a roadmap for applying SNA to future public health challenges. The COVID-19 era served as a case study for understanding public health behaviors and challenges during global crises. SNA has enabled the efficient tracking of information flow and analysis of community behaviors, helping to quantify the spread of misinformation through social media networks. SNA also allowed for real-time tracking of shifts in vaccine acceptance or resistance, particularly in response to new policies, side-effect reports, or misinformation. By targeting the right influential nodes and network interactions, public health campaigns can ensure that accurate information reaches communities faster and more effectively.

The World Health Organization's strategy for combating COVID-19 disinformation and vaccine hesitancy emphasizes the need for efforts to be "*community-led, data-driven, and collaborative*" [79]. Collaborative efforts between public health professionals and community

members foster trust, encourage rich community participation, and enhance cultural relevance and effectiveness of evidence-based programs. Such initiatives are critical for marginalized and underserved communities [80], allowing health professionals to track specific health data and intervention effectiveness, ultimately leading to more equitable outcomes.

In addition, SNA can provide insights into public health compliance, mapping community responses to guidelines on social distancing, mask usage, and lockdowns. By analyzing network interactions, researchers could trace how public health messages were received and shared, assessing community adherence or resistance to recommendations. Such insights are crucial for tailoring future interventions during health crises, ensuring they are effective and well targeted.

4.3. General Structural Characteristics of Networks and Analytical Methods

Most, though not all reviewed SNA studies (Tables 2–4), have applied standard network metrics or algorithms. Although advanced techniques in SNA have been developed to study phenomena such as information diffusion, dynamic evolution, and complex modeling strategies, these techniques are less frequently applied to social media data in public health informatics. One reason for this could be that these advanced methods often rely on detailed user behavior data related to actions and choices of individuals represented as nodes in the network. Without such data, it becomes difficult to apply models that depend on understanding the underlying actions and choices of network participants. While social media platforms provide large datasets, particularly in terms of user interactions, the availability of personal or ego-level information is limited. On platforms that promote anonymity, such as Reddit, there is almost no personal information available, further complicating the application of advanced SNA techniques.

Studies examining bot activity on Twitter have revealed a significant presence of non-humans posting on many controversial topics. This prevalence of bots poses challenges for public health studies, particularly those monitoring user opinions or behaviors. The inclusion of bot-generated content can distort research outcomes, especially as more researchers are applying machine learning algorithms to analyze large datasets, which may inadvertently obscure the presence of bot-made content, therefore skewing the results. The increasing number of bots spreading health misinformation on social media emphasizes the importance and urgency of developing improved verification and validation techniques that ensure the accuracy of public health research data derived from these platforms.

4.4. Ethical Use of Social Media Data

The public nature of social media data can often harm the privacy and anonymity of users, overshadowing consent in SNA research. This is especially critical when dealing with sensitive health-related topics that raise ethical considerations and a risk of reinforcing existing biases. This calls for a more comprehensive and inclusive approach in creating a regulatory framework for the ethics of such research while also including the community into the decision-making process for the downstream technical analyses. While this study was not scoped to develop or implement an ethics framework, other studies in the field have provided robust models to build on. For example, one study highlighted the importance of informed consent for all participants as well as respect for recruited participants and study communities [81]. Though these principles were developed in the context of human subjects for clinical trials, they can be evolved to be applicable to a modern era of public health research which includes utilizing large data samples from online repositories hosted on social media platforms.

With regard to informed consent, a central challenge and corresponding opportunity in the field is to glean to what degree researchers using web scraping for health research purposes should hold themselves to moral standards of traditional clinical research and explicitly engage (and enter into agreements) with content creators (i.e., people contributing to TikTok, X, and other social media platforms) about the collection and use of the

information they are gathering from them [82]. In terms of respect for communities, the authors [82] drew on the work of previous work [81] to define this as follows: *“clinical researchers can have obligations to research participants after they have been recruited and provided informed consent, and may also have obligations to the communities from which those participants come. For individuals, these can include confidentiality protections, the right to withdraw, compensation for study-related harms, continued post-research medical services, and dissemination of research results”*. In the modern era of utilizing social media data for public health research, these obligations could also include compensation for data use, options for being credited for data use, and being informed about the outputs from the research. Ultimately, evolving and implementing practices such as informed consent and respect for communities are foundational to building and maintaining ongoing trust with communities, and this principle applies to SNA research as well.

4.5. Limitations and Future Research

This review has several limitations. First, some pertinent studies were overlooked, particularly those utilizing Instagram (e.g., [83,84]) and Reddit (e.g., [85]) as platforms for harvesting social media data. These platforms were not as widely represented in the reviewed literature, which may be due to the restrictive nature of their public APIs. Although we followed the rigorous systematic review methodology based on PRISMA guidelines, not all relevant published literature were identified. This gap may have arisen due to omissions in the indexing of papers or databases as well as limitations with our search query. This review’s focus on widely available platforms may have unintentionally led to the exclusion of smaller or emerging platforms that are now gaining traction in the field of public health.

Furthermore, this review focuses solely on studies published in the English language and in the five-year period around the COVID-19 pandemic, which may limit the applicability of findings to other time periods. Because the pandemic dominated the types of SNA research in public health during the review period, other important public health topics may have been underexplored or overshadowed. Social media usage patterns and public health concerns can evolve rapidly, impacting the relevance of the results over time. Finally, the discussion of ethics related to SNA research conducted thus far may not be generalizable to the broader population due to the specific user demographics of different platforms and differing needs for ethical protections for certain marginalized populations [86].

Future research should investigate the trends in SNA research on social media data in the periods preceding and following the time period covered in this review. This would provide a more comprehensive understanding of the evolution of public health topics since the inception of social media platforms. This could be supported by considering the integration of SNA with other interdisciplinary methods and analytical frameworks that could provide a more holistic understanding of public health issues.

Finally, changes in social media platform policies, such as those affecting data accessibility and API usage, can influence the scope and feasibility of future research. Future research could incorporate data from platforms like Instagram, TikTok, and other emerging social networks, particularly as access to their data becomes more feasible. These platforms may offer valuable insights into younger or different demographic groups that are underrepresented in studies using Twitter and Facebook data.

5. Conclusions

Our goal in writing this systematic review was to identify the current trends in research applying network analysis methods to social media data in order to facilitate public health informatics. We found that, likely due to the time period that we reviewed, the majority of the studies reviewed have focused on discourse related to COVID-19 on Twitter/X, while other platforms and health conditions were sparsely represented in comparison. We outlined the commonly used analytical approaches, algorithms, and software used by SNA

researchers. Finally, we found that only 33% of reviewed papers explicitly considered the ethical implications of their research.

In the field of public health, the ethical use of social media data is crucial for maintaining trust within communities, protecting privacy, and ensuring accuracy of all health-related information and research outcomes. Accurate data are crucial for epidemiologists and public health practitioners to understand disease patterns, identify specific risk factors, and develop effective public health interventions. The use of social media data without proper ethical considerations may skew epidemiological findings as the usage of various social media platforms may be more active in certain communities/populations, leading to conclusions that do not accurately represent the broader population [34]. Therefore, we highlight the need for more comprehensive and ethically sound research practices as well as the importance of broadening the scope of future studies to encompass diverse populations, by including multiple platforms in the analysis. Addressing these gaps will further strengthen the integration of social media analytics into public health informatics, ultimately contributing to more effective and timely public health responses.

Author Contributions: Conceptualization, S.G., P.W., A.J. and P.P.; methodology, S.G., M.L., J.F. and Y.S.; writing—original draft preparation, S.G., M.L., J.F., Y.S., A.J., A.S. and K.-H.C.; writing—review and editing, S.G., A.J., Y.S., P.W., A.S., J.F. and P.P.; supervision, P.W., P.P. and K.-H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- Kaplan, A.M.; Haenlein, M. Users of the world, unite! The challenges and opportunities of Social Media. *Bus. Horizons* **2010**, *53*, 59–68. [CrossRef]
- Wong, B.J.D. Top Social Media Statistics and Trends of 2024. Forbes Advisor. Published 18 May 2023. Available online: <https://www.forbes.com/advisor/business/social-media-statistics/> (accessed on 22 August 2024).
- Mendoza-Herrera, K.; Valero-Morales, I.; Ocampo-Granados, M.E.; Reyes-Morales, H.; Arce-Amaré, F.; Barquera, S. An Overview of Social Media Use in the Field of Public Health Nutrition: Benefits, Scope, Limitations, and a Latin American Experience. *Prev. Chronic Dis.* **2020**, *17*, E76. [CrossRef]
- Suarez-Lledo, V.; Alvarez-Galvez, J. Prevalence of Health Misinformation on Social Media: Systematic Review. *J. Med. Internet Res.* **2021**, *23*, e17187. [CrossRef] [PubMed]
- van der Meer, T.G.L.A.; Jin, Y. Seeking Formula for Misinformation Treatment in Public Health Crises: The Effects of Corrective Information Type and Source. *Health Commun.* **2019**, *35*, 560–575. [CrossRef] [PubMed]
- Chou, W.-Y.S.; Oh, A.; Klein, W.M.P. Addressing Health-Related Misinformation on Social Media. *JAMA* **2018**, *320*, 2417–2418. [CrossRef] [PubMed]
- Wolfe, A.W. Social network analysis: Methods and applications. *Am. Ethnol.* **1997**, *24*, 219–220. [CrossRef]
- Borgatti, S.P.; Agneessens, F.; Johnson, J.C.; Everett, M.G. *Analyzing Social Networks*; SAGE Publications, Inc.: Thousand Oaks, CA, USA, 2024.
- Campbell, E.; Salathé, M. Complex social contagion makes networks more vulnerable to disease outbreaks. *Sci. Rep.* **2013**, *3*, 1905. [CrossRef]
- Mubamba, C.; Ramsay, G.; Abolnik, C.; Dautu, G.; Gummow, B. Combining value chain and social network analysis as a viable tool for informing targeted disease surveillance in the rural poultry sector of Zambia. *Transbound. Emerg. Dis.* **2018**, *65*, 1786–1796. [CrossRef]
- Di Battista, G. (Ed.) *Graph Drawing: Algorithms for the Visualization of Graphs*; Prentice Hall: Saddle River, NJ, USA, 1999.
- Jacomy, M.; Venturini, T.; Heymann, S.; Bastian, M. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE* **2014**, *9*, e98679. [CrossRef]
- Newman, M.E.J. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8577–8582. [CrossRef]
- Barabási, A.L.; Pósfai, M. *Network Science*; Cambridge University Press: Cambridge, UK, 2016.
- Freeman, L.C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1978**, *1*, 215–239. [CrossRef]
- Moukarzel, S.; Caduff, A.; Rehm, M.; del Fresno, M.; Pérez-Escamilla, R.; Daly, A.J. Breastfeeding Communication Strategies, Challenges and Opportunities in the Twitter-Verse: Perspectives of Influencers and Social Network Analysis. *Int. J. Environ. Res. Public Health* **2021**, *18*, 6181. [CrossRef] [PubMed]
- Liu, S.; Young, S.D.; Liu, S.; Young, S.D. A survey of social media data analysis for physical activity surveillance. *J. Forensic Leg. Med.* **2018**, *57*, 33–36. [CrossRef] [PubMed]

18. Fiesler, C.; Proferes, N. "Participant" perceptions of Twitter Research ethics. *Social Media + Soc.* **2018**, *4*, 205630511876336. [\[CrossRef\]](#)
19. Zimmer, M. "But the data is already public": On the ethics of research in Facebook. *Ethic-Inf. Technol.* **2010**, *12*, 313–325. [\[CrossRef\]](#)
20. Chu, K.-H.; Colditz, J.; Sidani, J.; Zimmer, M.; Primack, B. Re-evaluating standards of human subjects protection for sensitive health data in social media networks. *Soc. Netw.* **2021**, *67*, 41–46. [\[CrossRef\]](#)
21. Beigi, G.; Liu, H. A Survey on Privacy in Social Media. *ACM/IMS Trans. Data Sci.* **2020**, *1*, 1–38. [\[CrossRef\]](#)
22. Weller, K.; Kinder-Kurlanda, K. Uncovering the Challenges in Collection, Sharing and Documentation: The Hidden Data of Social Media Research? *Proc. Int. AAAI Conf. Web Soc. Media* **2021**, *9*, 28–37. [\[CrossRef\]](#)
23. Yoon, S.; Bakken, S. Methods of knowledge discovery in tweets. *NI* **2012**, *2012*, 463.
24. Hinds, J.; Williams, E.J.; Joinson, A.N. "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *Int. J. Human-Computer Stud.* **2020**, *143*, 102498. [\[CrossRef\]](#)
25. DeMuro, P.; Petersen, C. Legal and Regulatory Considerations Associated with Use of Patient-Generated Health Data from Social Media and Mobile Health (mHealth) Devices. *Appl. Clin. Inform.* **2015**, *6*, 16–26. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Chen, K.; Duan, Z.; Yang, S. Twitter as research data. *Politi-Life Sci.* **2021**, *41*, 114–130. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Yum, S. The COVID-19 Response in North America. *Disaster Med. Public Health Prep.* **2022**, *17*, 1–27. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Hung, M.; Lauren, E.; Hon, E.S.; Birmingham, W.C.; Xu, J.; Su, S.; Hon, S.D.; Park, J.; Dang, P.; Lipsky, M.S. Social Network Analysis of COVID-19 Sentiments: Application of Artificial Intelligence. *J. Med. Internet Res.* **2020**, *22*, e22590. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Xing, Y.; He, W.; Cao, G.; Li, Y. Using data mining to track the information spreading on social media about the COVID-19 outbreak. *Electron. Libr.* **2021**, *40*, 63–82. [\[CrossRef\]](#)
30. Eskandari, F.; Lake, A.A.; Butler, M. COVID-19 pandemic and food poverty conversations: Social network analysis of Twitter data. *Nutr. Bull.* **2022**, *47*, 93–105. [\[CrossRef\]](#)
31. Durazzi, F.; Müller, M.; Salathé, M.; Remondini, D. Clusters of science and health related Twitter users become more isolated during the COVID-19 pandemic. *Sci. Rep.* **2021**, *11*, 19655. [\[CrossRef\]](#)
32. Yum, S. Informatics for COVID-19 in New York and California. *Disaster Med. Public Health Prep.* **2021**, *16*, 1775–1779. [\[CrossRef\]](#)
33. Ahmed, W.; Vidal-Alaball, J.; Seguí, F.L.; Moreno-Sánchez, P.A. A Social Network Analysis of Tweets Related to Masks during the COVID-19 Pandemic. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8235. [\[CrossRef\]](#)
34. Pascual-Ferrá, P.; Alperstein, N.; Barnett, D.J. Social Network Analysis of COVID-19 Public Discourse on Twitter: Implications for Risk Communication. *Disaster Med. Public Health Prep.* **2020**, *16*, 561–569. [\[CrossRef\]](#)
35. Massaro, M.; Tamburro, P.; La Torre, M.; Mas, F.D.; Thomas, R.; Cobiauchi, L.; Barach, P. Non-pharmaceutical Interventions and the Infodemic on Twitter: Lessons Learned from Italy during the COVID-19 Pandemic. *J. Med. Syst.* **2021**, *45*, 50. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Nazar, S.; Pieters, T. Plandemic Revisited: A Product of Planned Disinformation Amplifying the COVID-19 "infodemic". *Front. Public Health* **2021**, *9*, 649930. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Weng, Z.; Lin, A. Public Opinion Manipulation on Social Media: Social Network Analysis of Twitter Bots during the COVID-19 Pandemic. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16376. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Ahmed, W.; Vidal-Alaball, J.; Downing, J.; Seguí, F.L. COVID-19 and the 5G Conspiracy Theory: Social Network Analysis of Twitter Data. *J. Med. Internet Res.* **2020**, *22*, e19458. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Gruz, A.; Mai, P. Going viral: How a single tweet spawned a COVID-19 conspiracy theory on Twitter. *Big Data Soc.* **2020**, *7*, 205395172093840. [\[CrossRef\]](#)
40. Kothari, A.; Walker, K.; Burns, K. #CoronaVirus and public health: The role of social media in sharing health information. *Online Inf. Rev.* **2022**, *46*, 1293–1312. [\[CrossRef\]](#)
41. Alieva, I.; Robertson, D.; Carley, K.M. Localizing COVID-19 Misinformation: A Case Study of Tracking Twitter Pandemic Narratives in Pennsylvania Using Computational Network Science. *J. Health Commun.* **2023**, *28*, 76–85. [\[CrossRef\]](#)
42. Duzen, Z.; Riveni, M.; Aktas, M.S. Analyzing the Spread of Misinformation on Social Networks: A Process and Software Architecture for Detection and Analysis. *Computers* **2023**, *12*, 232. [\[CrossRef\]](#)
43. Ahmed, W.; Seguí, F.L.; Vidal-Alaball, J.; Katz, M.S. COVID-19 and the "Film Your Hospital" Conspiracy Theory: Social Network Analysis of Twitter Data. *J. Med. Internet Res.* **2020**, *22*, e22374. [\[CrossRef\]](#)
44. Seckin, O.C.; Atalay, A.; Otenen, E.; Duygu, U.; Varol, O. Mechanisms driving online vaccine debate during the COVID-19 pandemic. *Social Media + Soc.* **2024**, *10*. [\[CrossRef\]](#)
45. Durmaz, N.; Hengirmen, E. The dramatic increase in anti-vaccine discourses during the COVID-19 pandemic: A social network analysis of Twitter. *Hum. Vaccines Immunother.* **2022**, *18*, 2025008. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Boucher, J.-C.; Cornelson, K.; Benham, J.L.; Fullerton, M.M.; Tang, T.; Constantinescu, C.; Mourali, M.; Oxoby, R.J.; A Marshall, D.; Hemmati, H.; et al. Analyzing Social Media to Explore the Attitudes and Behaviors Following the Announcement of Successful COVID-19 Vaccine Trials: Infodemiology Study. *JMIR Infodemiology* **2021**, *1*, e28800. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Villegas, C.; Ortiz, A.; Arriagada, V.; Ortega, S.; Walker, J.; Arriagada, E.; Kalergis, A.M.; Huepe, C. Influence of online opinions and interactions on the Covid-19 vaccination in Chile. *Sci. Rep.* **2022**, *12*, 21288. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Hobbs, A.; Aldosery, A.; Kostkova, P. Low credibility URL sharing on Twitter during reporting linking rare blood clots with the Oxford/AstraZeneca COVID-19 vaccine. *PLOS ONE* **2024**, *19*, e0296444. [\[CrossRef\]](#) [\[PubMed\]](#)

49. Ahmed, W.; Das, R.; Vidal-Alaball, J.; Hardey, M.; Fuster-Casanovas, A. Twitter's Role in Combating the Magnetic Vaccine Conspiracy Theory: Social Network Analysis of Tweets. *J. Med. Internet Res.* **2023**, *25*, e43497. [\[CrossRef\]](#)
50. Feliciano, J.T.; Salmi, L.; Blotner, C.; Hayden, A.; Nduom, E.K.; Kwan, B.M.; Katz, M.S.; Claus, E.B. Brain Tumor Discussions on Twitter (#BTSM): Social Network Analysis. *J. Med. Internet Res.* **2020**, *22*, e22005. [\[CrossRef\]](#)
51. Zhang, Z.; Ahmed, W. A comparison of information sharing behaviours across 379 health conditions on Twitter. *Int. J. Public Health* **2018**, *64*, 431–440. [\[CrossRef\]](#)
52. Liang, H.; Fung, I.C.-H.; Tse, Z.T.H.; Yin, J.; Chan, C.-H.; Pechta, L.E.; Smith, B.J.; Marquez-Lameda, R.D.; Meltzer, M.I.; Lubell, K.M.; et al. How did Ebola information spread on twitter: Broadcasting or viral spreading? *BMC Public Health* **2019**, *19*, 438. [\[CrossRef\]](#)
53. Aktuna, G.; Bahar-Özvarış, Ş. Investigating the aftermath of the Türkiye 2023 earthquake: Exploring post-disaster uncertainty among Syrian migrants using social network analysis with public health approach. *Front. Public Health* **2023**, *11*, 1204589. [\[CrossRef\]](#)
54. Boucher, J.-C.; Kim, S.Y.; Jessiman-Perreault, G.; Edwards, J.; Smith, H.; Frenette, N.; Badami, A.; Scott, L.A. HPV vaccine narratives on Twitter during the COVID-19 pandemic: A social network, thematic, and sentiment analysis. *BMC Public Health* **2023**, *23*, 694. [\[CrossRef\]](#)
55. Moukarzel, S.; Rehm, M.; Caduff, A.; del Fresno, M.; Perez-Escamilla, R.; Daly, A.J. Real-time Twitter interactions during World Breastfeeding Week: A case study and social network analysis. *PLOS ONE* **2021**, *16*, e0249302. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Goyal, K.; Nafri, A.; Marwah, M.; Aramadaka, S.; Aggarwal, P.; Malhotra, S.; Mannam, R.; Gupta, O.; Malhotra, K. Evaluating the Global Impact of Stroke Awareness Month: A Serial Cross-Sectional Analysis. *Cureus* **2022**, *14*, e28997. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Pratiwi, H.; Benkő, R.; Kusuma, I.Y. Navigating the asthma network on Twitter: Insights from social network and sentiment analysis. *Digit. Health* **2024**, *10*. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Martínez-Martínez, F.; Roldán-Álvarez, D.; Martín, E.; Hoppe, H.U. An analytics approach to health and healthcare in citizen science communications on Twitter. *Digit. Health* **2023**, *9*, 205520762211453. [\[CrossRef\]](#)
59. Gomaa, B.T.; Walsh-Buhi, E.R.; Funk, R.J. Understanding Melanoma Talk on Twitter: The Lessons Learned and Missed Opportunities. *Int. J. Environ. Res. Public Health* **2022**, *19*, 11284. [\[CrossRef\]](#)
60. Park, A. Tweets Related to Motivation and Physical Activity for Obesity-Related Behavior Change: Descriptive Analysis. *J. Med. Internet Res.* **2022**, *24*, e15055. [\[CrossRef\]](#)
61. Cai, M.; Luo, H.; Meng, X.; Cui, Y. Differences in Behavioral Characteristics and Diffusion Mechanisms: A Comparative Analysis Based on Social Bots and Human Users. *Front. Phys.* **2022**, *10*, 875574. [\[CrossRef\]](#)
62. Lv, W.; Zhou, W.; Gao, B.; Han, Y.; Fang, H. New Insights Into the Social Rumor Characteristics During the COVID-19 Pandemic in China. *Front. Public Health* **2022**, *10*, 864955. [\[CrossRef\]](#)
63. Sun, K.; Wang, H.; Zhang, J. The impact factors of social media users' forwarding behavior of COVID-19 vaccine topic: Based on empirical analysis of Chinese Weibo users. *Front. Public Health* **2022**, *10*, 871722. [\[CrossRef\]](#)
64. Wang, D.; Zhou, Y.; Ma, F. Opinion Leaders and Structural Hole Spanners Influencing Echo Chambers in Discussions About COVID-19 Vaccines on Social Media in China: Network Analysis. *J. Med. Internet Res.* **2022**, *24*, e40701. [\[CrossRef\]](#)
65. Cai, M.; Luo, H.; Meng, X.; Cui, Y. Topic-Emotion Propagation Mechanism of Public Emergencies in Social Networks. *Sensors* **2021**, *21*, 4516. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Shi, J.; Khoo, Z. Online health community for change: Analysis of self-disclosure and social networks of users with depression. *Front. Psychol.* **2023**, *14*, 1092884. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Li, S.; Liu, Z.; Li, Y. Temporal and spatial evolution of online public sentiment on emergencies. *Inf. Process. Manag.* **2020**, *57*, 102177. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Hoffman, B.L.; Felter, E.M.; Chu, K.-H.; Shensa, A.; Hermann, C.; Wolynn, T.; Williams, D.; Primack, B.A. It's not all about autism: The emerging landscape of anti-vaccination sentiment on Facebook. *Vaccine* **2019**, *37*, 2216–2223. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Deng, W.; Yang, Y. Cross-Platform Comparative Study of Public Concern on Social Media during the COVID-19 Pandemic: An Empirical Study Based on Twitter and Weibo. *Int. J. Environ. Res. Public Health* **2021**, *18*, 6487. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Kwon, S.; Park, A. Examining thematic and emotional differences across Twitter, Reddit, and YouTube: The case of COVID-19 vaccine side effects. *Comput. Hum. Behav.* **2023**, *144*, 107734. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Gruzdz, A.; Abul-Fottouh, D.; Song, M.Y.; Saiphoo, A. From Facebook to YouTube: The potential exposure to COVID-19 Anti-Vaccine videos on social media. *Social Media + Soc.* **2023**, *9*, 205630512211504. [\[CrossRef\]](#)
72. Feng, Y.; Li, X.; Ma, X.; Zhu, Z.; Chen, K.; Gao, J.; Xia, J.; Jiang, R.; Lu, J. Using online social networks to provide a parental health-education intervention for preventing unintentional injuries among children aged 0–3 years: A randomized controlled trial and social network analysis in Shanghai, China. *Front. Public Health* **2023**, *10*, 1049851. [\[CrossRef\]](#)
73. Yao, Z.; Ni, Z.; Zhang, B.; Du, J. Do Informational and Emotional Elements Differ between Online Psychological and Physiological Disease Communities in China? A Comparative Study of Depression and Diabetes. *Int. J. Environ. Res. Public Health* **2022**, *19*, 2167. [\[CrossRef\]](#)
74. Fu, L.; Sun, H.; Xu, K. Whether Social Participation Can Affect the Central Government Public Policy Response to the COVID-19 in China. *Front. Public Health* **2022**, *10*, 842373. [\[CrossRef\]](#)
75. Zhan, Q.; Zhuo, W.; Liu, Y. Social Influence Maximization for Public Health Campaigns. *IEEE Access* **2019**, *7*, 151252–151260. [\[CrossRef\]](#)

76. Ahmed, W.; Lugovic, S. Social media analytics: Analysis and visualisation of news diffusion using NodeXL. *Online Inf. Rev.* **2019**, *43*, 149–160. [CrossRef]
77. Wang, D.; Qian, Y. Echo Chamber Effect in Rumor Rebuttal Discussions About COVID-19 in China: Social Media Content and Network Analysis Study. *J. Med. Internet Res.* **2021**, *23*, e27009. [CrossRef] [PubMed]
78. Social Media Fact Sheet. Pew Research Center. Published 13 January 2024. Available online: <https://www.pewresearch.org/internet/fact-sheet/social-media/> (accessed on 8 July 2024).
79. World Health Organization. *COVID-19 Global Risk Communication and Community Engagement Strategy, December 2020–May 2021: Interim Guidance, 23 December 2020*; World Health Organization: Geneva, Switzerland, 2020.
80. Chung-Do, J.J.; Bifulco, K.; Antonio, M.; Tydingco, T.; Helm, S.; Goebert, D.A. A cultural analysis of the NAMI-NH Connect Suicide Prevention Program by rural community leaders in Hawai'i. *J. Rural. Ment. Health* **2016**, *40*, 87–102. [CrossRef]
81. Emanuel, E.J.; Wendler, D.; Grady, C. What Makes Clinical Research Ethical? *JAMA* **2000**, *283*, 2701–2711. [CrossRef] [PubMed]
82. Rennie, S.; Buchbinder, M.; Juengst, E.; Brinkley-Rubinstein, L.; Blue, C.; Rosen, D.L. Scraping the web for public health gains: Ethical considerations from a 'Big Data' research project on HIV and incarceration. *Public Health Ethics* **2020**, *13*, 111–121. [CrossRef]
83. Vassey, J.; Valente, T.; Barker, J.; Stanton, C.; Li, D.; Laestadius, L.; Cruz, T.B.; Unger, J.B. E-cigarette brands and social media influencers on Instagram: A social network analysis. *Tob. Control.* **2022**, *32*, e184–e191. [CrossRef]
84. Massey, P.M.; Kearney, M.D.; Hauer, M.K.; Selvan, P.; Koku, E.; E Leader, A. Dimensions of Misinformation About the HPV Vaccine on Instagram: Content and Network Analysis of Social Media Characteristics. *J. Med. Internet Res.* **2020**, *22*, e21451. [CrossRef]
85. Moessner, M.; Feldhege, J.; Wolf, M.; Bauer, S. Analyzing big data in social media: Text and network analyses of an eating disorder forum. *Int. J. Eat. Disord.* **2018**, *51*, 656–667. [CrossRef]
86. Jaiswal, A.; Shah, A.; Harjadi, C.; Windgassen, E.; Washington, P. Ethics of the Use of Social Media as Training Data for Artificial Intelligence Models used for Digital Phenotyping: Commentary (Preprint). *JMIR Form. Res.* **2024**, *8*, e59794. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.