

DOI: [https://doi.org/10.48009/1\\_iis\\_2021\\_73-85](https://doi.org/10.48009/1_iis_2021_73-85)

## **Proliferation of health misinformation on social media platforms: a systematic literature review**

**Piyush Vyas**, *Dakota State University*, [piyush.vyas@trojans.dsu.edu](mailto:piyush.vyas@trojans.dsu.edu)

**Gitika Vyas**, *Rajasthan Technical University*, [gitika.dadhich@gmail.com](mailto:gitika.dadhich@gmail.com)

**Jun Liu**, *Dakota State University*, [jun.liu@dsu.edu](mailto:jun.liu@dsu.edu)

### **Abstract**

This paper presents a systematic literature review to provide a conceptual framework and future research directions for the identification and detection of health misinformation on social networking sites. We have found the health misinformation problem is yet to evolve and requires more exploration towards the identification of reliable techniques. We have extracted the existing literature from five research databases. The focus of this review is to explore various concepts, techniques, and datasets adopted by the prior studies. After applying Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines we have got 27 final papers to include in this review. The findings show that despite adopting various techniques there was a lack of research done on concepts like fake news and healthcare fact-checking. Furthermore, this paper also identified significant feature and technique-based research gaps for future research directions.

**Keywords:** health misinformation, literature review, social media, fake news, fact-checking.

### **Introduction**

Due to global digitization and inexpensive Internet services, various web portals and websites became encyclopedias for healthcare-related information. More than 4 billion unique users joined the realm of the internet and the numbers are increasing day by day (Barua et al., 2020). It has been observed that more than 70% of adults utilized various web sources to search healthcare-related information (Barua et al., 2020). Among various Internet services, social media platforms are used as the main source of information by people whereby they interact with each other to create, share, and exchange information among their virtual communities (Niknam et al., 2020). Facebook, Twitter, and Instagram are the social media giants. An unrestrained usage and access of social media platforms made them a discussion portal and healthcare are an emerging topic among several other issues (Ghenai & Mejova, 2018). Before consulting a physician, 44% of people often search about the cure of an illness on the internet and change their decision about the treatment (Ghenai, 2017). However, the quality of information and content posting intentions were always been questioned and must be treated as a potential threat to society (Ghenai & Mejova, 2018). Moreover, the influence of propagated healthcare misinformation on people often leads them to harmful health decisions (Ghenai, 2017).

Therefore, usage of social media and people's reliance on available content originates a problem of health misinformation dissemination. Misinformation is the piece of content that fulfills the deceiving intention by spreading misleading information. Misinformation can be categorized in many forms such as hoaxes, rumors, fake news, fake reviews, and false facts (Vyas & El-Gayar, 2020). These categories keep on increasing. For example, in the current pandemic situation, "Infodemic" emerged as a new category that

comprises the usage of unreliable information, especially for healthcare. Hence, healthcare professionals are often facing the problem to eliminate non-credible information from all official and non-official sources (Datta et al., 2020). Furthermore, in recent years, the proliferation of health misinformation gauged the attention of researchers. Abul-Fottouh et al. (2020), Barua et al. (2020), Datta et al. (2020), Ghenai & Mejova (2018), Hou et al. (2019), and Niknam et al. (2020) studied the propagation of health misinformation on social media platforms. Lavorgna et al. (2018), Oberiri and Omar (2020), Secosan et al. (2020), and Shi et al. (2019) delved into spread of health fake news, and Boudewyns et al. (2020), Cui and Lee (2020), and Kawchuk et al. (2020) explored about the dispersion of false facts on fact-checking websites (i.e., Politifact.com & Snopes.com) and social media. Hence, the focus of this paper is to provide a systematic and comprehensive literature review of all relevant literature on the proliferation of health misinformation on social media.

Accordingly, this paper leverages existing research papers on the topic of healthcare, misinformation, and social media with the viewpoint of studying various concepts, techniques, and datasets available in the quest of resolving the health misinformation problem on social media. Thus, the following research questions have emerged from our focus:

*RQ1. What are the different concepts(themes) for health misinformation on social media studied in existing literature?*

*RQ2. What are the techniques utilized to address/overcome/detect health misinformation on social media sites?*

*RQ3. What are the datasets available to develop the detection and identification systems for rectifying the problem of health misinformation?*

The paper is structured as follows. Section 2 presents the background. Section 3 posits the methodology followed by literature analysis to explore the conceptual framework in section 4. Section 5 provides the discussion and explores the answers to identified research questions, future research questions, and section 6 concludes our research.

## Background

Social media as a wide-open gateway to public exposure have been growing across the globe with variety of textual information in the form of blogs, microblogs and short texts (Koohang et al., 2018; Mađra-Sawicka et al., 2020; Wong et al., 2020). The understanding of such information is critical especially when the information is related to healthcare because it can have enormous impact on public health (Troshani & Wickramasinghe, 2018). Lately, researchers have targeted health misinformation on social media platforms as a potential problem. Their different approaches often encircled around finding the effects of misleading content on societal perception. Studies such as (Abul-Fottouh et al., 2020; Kouzy et al., 2020; Oberiri & Omar, 2020; Porat et al., 2019; Shi et al., 2019; Smaldone et al., 2020; Waszak et al., 2018) considered the statistical analysis approach as a tool to explore the effects of misinformation, whereas studies such as (Cui et al., 2020; Cui & Lee, 2020; Ghenai, 2017; Ghenai & Mejova, 2018; Glasdam & Stjernswärd, 2020; Hou et al., 2019; Islam et al., 2020; J. Li et al., 2020; Niknam et al., 2020; Sear et al., 2020; Silva et al., 2020) focused on the identification of different factors or features to detect and eliminate the misinformation by utilizing state of the art machine learning (ML) and deep learning (DL) techniques.

Machine learning techniques offer the opportunity to discover hidden patterns and relationships in the data (Christozov & Mitreva, 2020). Therefore, as a potential problem, researchers have adopted the classification of user-generated content as health misinformation or genuine content. Niknam et al. (2020) performed a thematic analysis on Instagram posts by extracting user-search tags (i.e., upload date, media type) to

characterize the public health information as misinformation. Ghenai & Mejova (2018) classified the misinformation related to fake cancer cures on Twitter by utilizing the logistic regression technique. They have extracted user-centric, tweet-centric, readability-centric, and medical-centric features. (See table 1). Hou et al. (2019) adopted viewer-centric and linguistic features to automatically detect prostate cancer-related misinformation on YouTube by utilizing the support vector machine classification technique. Silva et al. (2020) has performed a classification task by utilizing machine learning techniques to investigate the relationship between users and misinformation through tweet-related, engagement related and user-related features. Sear et al. (2020) performed the topic modeling to extract the frequent health misinformation-related themes in Facebook posts by Latent Dirichlet allocation (LDA) techniques through calculating coherence and subjectivity scores. J. Li et al. (2020) utilized the ML-based exploratory analysis to classify the health-related false facts on Snopes.com and Poynter.com using fact-checking reviews, source content, social engagement, and user reply features. Cui et al. (2020) adopted the linguistic features to detect health fake news through graph attention neural network. Further, Cui et al. (2020) used a gated recurrent unit (GRU) to detect health misinformation on social media platforms by using user engagement features.

Furthermore, Abul-Fottouh et al. (2020) observed two interconnected concepts- Homophily and Filter-bubble effects over the vaccination-based YouTube videos by utilizing a Latent Order Logistic Model (LOLOG). Oberiri & Omar (2020) did an online survey implementing gratification theory on WhatsApp messages and Facebook posts to predict COVID-19 fake news by partial least square-based structure equation modeling (PLS-SEM) technique. Waszak et al. (2018) analyzed the misinformation and falsehood of health news on social media platforms like Twitter, Facebook, Pinterest, LinkedIn with the help of analytical tool BuzzSumo. Smaldone et al. (2020) surveyed user-centric data on Facebook, Twitter, Instagram, YouTube, Pinterest of both qualitative and quantitative nature by Structural Equation Model (SEM). Shi et al. (2019) scrutinized online false news of a cannabis cancer cure and its impact using Google Trends' relative search volume (RSV) tool. Porat et al. (2019) statistically analyzed the diphtheria-related tweets by applying Pearson Correlation for content and source authenticity. Kouzy et al. (2020) evaluated the enormity of propagated coronavirus misinformative tweets statistically by employing the Chi-square method.

## Methodology

To structure our systematic literature review, we used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines proposed by Liberati et al. (2009). These guidelines offer a checklist of 27 items under four-phase flow diagrams for establishing transparency in the process of systematic review.

### Data Sources and Search Strategy

This work has used five databases- ABI/Inform, ACM Digital Library, PubMed, Science Direct, and Web of science- to extract the most relevant literature presented in the English language. The search query is encircled around the three main keywords- "Healthcare" AND "Misinformation" AND "Social Media" to capture the studies during the period of five years – from 2015 to 2020.

### Eligibility Criteria

#### *Inclusion criteria*

1. Studies should have discussed health misinformation with respect to social media platforms.
2. The focus or aim of studies should be clear and targets the proliferation/detection of health misinformation.

3. The studies should utilize adequate methodologies/techniques to target the problem of health misinformation.
4. The datasets utilized should strictly align with the aim of the studies.

## Exclusion criteria

All of those studies were excluded that performed literature reviews or meta-analyses or did not discuss misinformation proliferation in the context of healthcare and social media.

## Evaluation Criteria

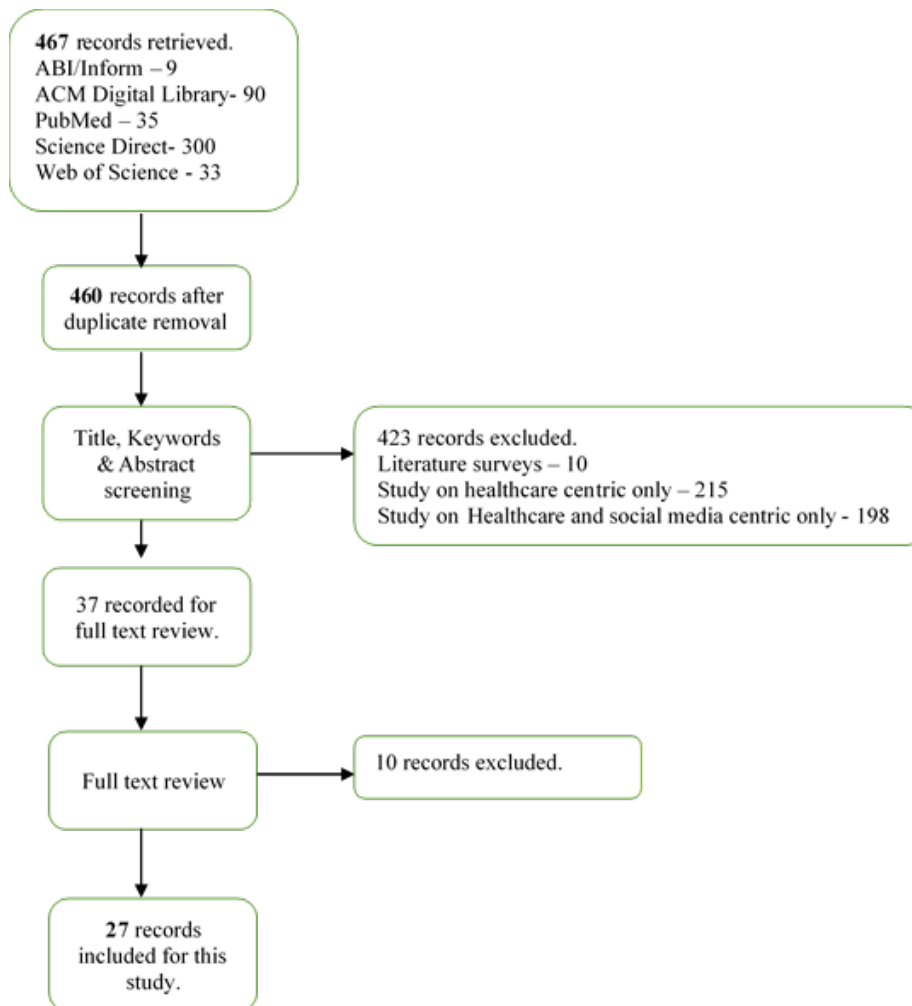


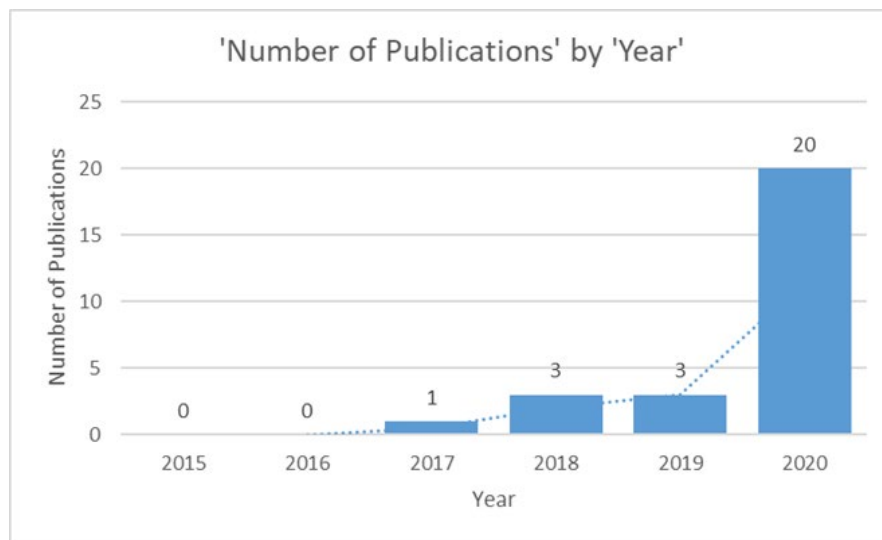
Figure 1. PRISMA workflow of this survey

This work has utilized PRISMA methodology and conceptual framework as a study evaluation criterion for selected literature. First, we reviewed the title and keywords of the articles. Second, a review of abstracts was conducted. Third, a reading of the full text was conducted to extract the answers regarding identified research questions in the context of various techniques and concepts used to target health misinformation.

## Study Selection

Figure 1 shows all phases according to PRISMA guidelines. This includes study identification, screening, eligibility assessment, and included articles. As illustrated, a total of 467 records were extracted by initial search query from five datasets- ABI/Inform (9), ACM Digital Library (90), PubMed (35), Science Direct (300), and Web of Science (33). After an assessment of duplicate articles, we got 460 pieces of literature for further filtering. A review of titles, keywords, and abstracts has been conducted to exclude/include the literature that consists of content of misinformation on social media in the healthcare domain. As a result, 423 studies were excluded which were - literature surveys (10), healthcare-related context (215) without focusing on social media or misinformation, and healthcare-related studies on social media without any focus on the proliferation of misinformation (198). 37 of the records were considered for full-text reading and 10 of them were excluded because they were strictly theoretical in nature. Finally, 27 records were selected for this systematic review that includes all required fields – the healthcare domain, the proliferation of misinformation, and social media platforms. To select the final records, the focus was on utilized study type (Ex., Survey), techniques (Ex., Classification), and tools (Ex., Google trends).

## Research Trends



**Figure 2. Research Trends by selecting studies.**

Figure 2 shows the research trends associated with the proliferation of health misinformation on social media platforms. As we considered the last five years' literature for this systematic review, we found that, in the first two years including 2015 and 2016, there was not any relevant literature aligned with our focus. Furthermore, most of the relevant literature belongs to the year 2020, the reason behind this is a dispersion of COVID-19 “Infodemic” on social media. Overall, there is a random pattern in publication trends however the moving average indicates the surge in coming years.

## Exploring the Research Questions

With respect to our conceptual framework, to answer the research questions, various concepts used in the literature were explored. Some of the literature discussed more than one concept (i.e., misinformation and fake news altogether). In general, 16 of them targeted the problem of health misinformation, 3 of them specifically discussed the healthcare fact-checking problem, and 7 of them explored the dissemination of healthcare fake news on social media.

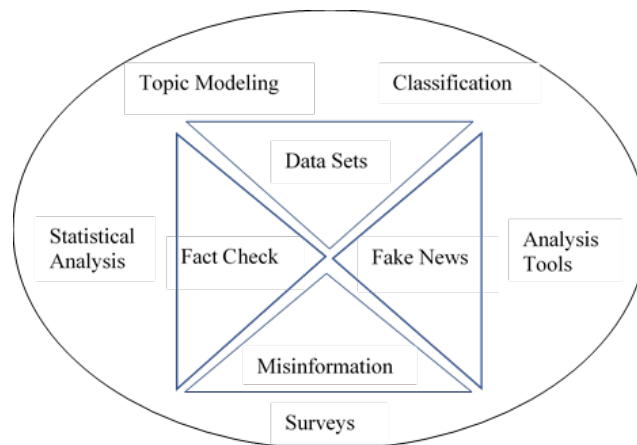
Literature Analysis

Table 1. Distribution of articles by concepts, platforms, categories, and techniques

| Concepts       | Platforms  | Categories                              | Techniques   | Author                        |
|----------------|--|---|--|-------------------------------|
| Misinformation | Instagram  | Topic modeling                          | customized input-output quantitative (IPO)                         | (Niknam et al., 2020)         |
|                | Government websites                              | Survey/statistical analysis             | Chi-square test / STATA tool                                       | (Datta et al., 2020)          |
|                | All social media                                 | Survey/statistical analysis             | Structural equation Modeling (SEM)                                 | (Barua et al., 2020)          |
|                | YouTube  | Statistical analysis                    | Latent Order Logistic Model (LOLOG)                                | (Abul-Fottouh et al., 2020)   |
|                | Twitter  | -                                       | -  | (Venegas-Vera et al., 2020)   |
|                | Twitter  | Classification                          | Logistic regression  | (Ghenai & Mejova, 2018)       |
|                | Twitter  | Classification                          | Statistical analysis/ Tools (botornot, metamap, Health on the Net) | (Ghenai, 2017)                |
|                | YouTube  | Classification                          | Support Vector Classifier  | (Hou et al., 2019)            |
|                | Twitter, Facebook, Pinterest, LinkedIn           | Statistical Analysis                    | Buzzsumo   | (Waszak et al., 2018)         |
|                | Facebook, Twitter, Instagram, YouTube, Pinterest | Survey / exploratory study              | SEM  | (Smaldone et al., 2020)       |
|                | Twitter  | Classification                          | Machine Learning (ML)  | (Silva et al., 2020)          |
|                | Facebook   | Topic modeling                          | LDA  | (Sear et al., 2020)           |
|                | Twitter  | Statistical analysis                    | Pearson correlation  | (Porat et al., 2019)          |
|                | Twitter  | Statistical analysis                    | Chi-square statistic   | (Kouzy et al., 2020)          |
| Fact checks    | Questionnaires on social media                   | Topic modeling /survey                  | Manual approach  | (Glasdam & Stjernswärd, 2020) |
|                | Blogs drug prescription                          | Survey/Statistical Analysis             | Regression / SPSS tool   | (Boudewyns et al., 2020)      |
| Fake news      | -  | Analysis/Market Review Tool, MRT        | -  | (Kawchuk et al., 2020)        |
|                | WhatsApp / Facebook                              | Online survey                           | PLS-SEM  | (Oberiri & Omar, 2020)        |
|                | Google   | Google Trends' statistical analysis     | Relative search volume tool  | (Shi et al., 2019)            |
|                | Smsocialnetwork.com                              | Pilot study                             | -  | (Lavorgna et al., 2018)       |
|                | Facebook   | Survey on Webropol.                     | PLS-SEM & Neural Network /SPSS tool                                | (Islam et al., 2020)          |
| Dataset        | Knowlife.mpi-inf.mpg.de                          | Classification / Knowledge graph        | Graph Neural Net (GNN) deterrent                                   | (Cui et al., 2020)            |
|                | Twitter  | Classification / dataset                | ML Techniques and GRU  | (Cui & Lee, 2020)             |
|                | Twitter/ Reuters and public radio                | Dataset                                 | -  | (Zhou et al., 2020)           |
|                | Twitter  | Statistical analysis/dataset            | Descriptive statistics   | (Lopez & Gallemore, 2020)     |
|                | Politifact.com, snopes.com                       | Exploratory analysis and classification | ML Techniques  | (Y. Li et al., 2020)          |

Table 1 shows the distribution of selected literature by utilized concepts, techniques, social media platforms, and categories along with the references. Although few studies did not comprise all table parameters. Furthermore, the discussion section explains these concepts by addressing our identified research questions.

Figure 3 shows our proposed conceptual framework that comprises various identified concepts such as fact-checking, fake news, misinformation, and datasets. The reason being extracting these research terms as concepts from the existing literature is to provide overall knowledge about numerous research directions and availability of datasets to current researchers. These concepts were highlighted and considered by studied literature such as, (Abul-Fottouh et al., 2020; Barua et al., 2020; Boudewyns et al., 2020; Cui & Lee, 2020; Datta et al., 2020; Ghenai & Mejova, 2018; Hou et al., 2019; Lavorgna et al., 2018; Niknam et al., 2020; Oberiri & Omar, 2020; Secosan et al., 2020).



**Figure 3. Conceptual Framework for Health Misinformation on social media.**

The inner shell of the conceptual framework sheds the light on aforementioned health misinformation categories. Most of the existing literature such as (Abul-Fottouh et al., 2020; Barua et al., 2020; Datta et al., 2020; Ghenai & Mejova, 2018; Hou et al., 2019; Niknam et al., 2020) considered “misinformation” as a generalized term rather than specifying in a particular category (i.e., fake news, fake reports, false facts). Therefore, we have considered misinformation as a separate concept; otherwise, significant literature would have left out from this conceptual framework. The outermost layer of this conceptual framework comprises various techniques such as topic modeling, statistical analysis, quantitative/qualitative surveys, analytical tools, and classification algorithms adopted by Barua et al. (2020), Hou et al. (2019), Kawchuk et al. (2020), Niknam et al. (2020), and Sear et al. (2020).

## Discussion

Almost all social media platforms such as YouTube, Facebook, Instagram, WhatsApp, and Twitter possess various signals in the form of user-generated content to identify health misinformation. Often these signals are treated as potential features to detect misleading content either as a user comment, posts, tweets, and audio/video share. Hence, prior studies posit user-centric, text-centric, post-centric, and linguistic-centric feature sets. The following subsections answer the identified research questions and explore the proposed conceptual framework.

## **RQ1. Different concepts(themes) for health misinformation on social media**

Silva et al. (2020) state that the proliferation of misinformation is not contemporary but persists from the time of the cold war. However, we are observing this now, in the era of social media. Misinformation can be any form of social media content (textual, visual, and audio) with misleading intentions (Venegas-Vera et al., 2020). Fact-checking is a way of identifying reliable information with significant evidence. Politifact.com and Snopes.com are two major platforms for fact-checking (Venegas-Vera et al., 2020). Fake news comprises deceiving content with false evidence, hence this is the biggest threat to society. People shares their health issues and treatment experiences on social networking sites for seeking help and suggestions. Therefore social networking sites can be allies and enemies of society at the same time (Venegas-Vera et al., 2020).

Moreover, healthcare fake news is a multifaced problem with various categories such as Fabricated news, Manipulated news, Advertisement news, Irrelevant news, and Sufficient news (Waszak et al., 2018). However, fact-checks provide a reliable way to identify timely pieces of misinformation, fact-checkers cannot address every piece of misinformation and their professional work necessarily involves various selection biases as they focus on scarce resources (Venegas-Vera et al., 2020). Furthermore, the availability of the datasets that comprise misleading information would help researchers and practitioners to understand the patterns and behavior of such content so that they can further use extracted knowledge for the elimination and detection of health misinformation.

We have found that Abul-Fottouh et al. (2020), Barua et al. (2020), Datta et al. (2020), Ghenai (2017), Ghenai and Mejova (2018), Hou et al. (2019), Lopez and Gallemore (2020), Niknam et al. (2020), Porat et al. (2019), Sear et al. (2020), Silva et al. (2020), Smaldone et al. (2020), Venegas-Vera et al. (2020), and Waszak et al. (2018) have used the concept of misinformation. Cui et al. (2020), Islam et al. (2020), Lavorgna et al. (2018), Li et al. (2020), Secosan et al. (2020), and Shi et al. (2019) have used the concept of fake news. Boudewyns et al. (2020), Cui and Lee (2020), and Kawchuk et al. (2020) have used the concept of fact-checking. Furthermore, Cui and Lee (2020), Li et al. (2020), Lopez and Gallemore (2020), and Zhao et al. (2020) provides the different datasets for health misinformation on social media that are further described in Table 2.

## **RQ2. Techniques utilized to address the health misinformation on social media sites.**

Societal reliance on dispersed health misinformation on social media may lead to harmful events on a large scale. For example, it has been observed that the false claims regarding cannabis as a cancer cure would have engaged 4.26 million unique people on social media, whereas only 0.036 million people engaged in real news that debunks such claims (Shi et al., 2019). Therefore, advanced techniques such as natural language processing, machine learning, and deep learning approaches should be applied to eliminate health misinformation from social media platforms (Tasnim et al., 2020). In this review, we found out different techniques adopted by researchers, such as Glasdam and Stjernswärd (2020), Niknam et al. (2020), and Sear et al. (2020) have adopted the topic modeling (i.e., identification of various frequent themes of misinformation). Abul-Fottouh et al. (2020), Barua et al. (2020), and Boudewyns et al. (2020) have used statistical analysis (i.e., causes and effect of misinformation). Cui et al. (2020), Ghenai (2017), Ghenai and Mejova (2018), Hou et al. (2019), and Silva et al. (2020) have utilized machine learning and neural network classification techniques (i.e., supervised ML techniques such as logistic regression, support vector machine, graph neural networks). Datta et al. (2020), Islam et al. (2020), and Kawchuk et al. (2020) have used analytical tools (i.e., google trends, MRT, SPSS). Barua et al. (2020), Islam et al. (2020), and Smaldone et al. (2020) have used qualitative surveys (i.e., interviewing people). Furthermore, “these practices should be controlled with regulatory and law enforcement measures alongside ensuring telemedicine-based services providing accurate information on COVID-19 (Tasnim et al., 2020)”.



**RQ3. The datasets available**

Table 2 shows the available updated datasets along with their description and sources.

**Table 2. Various available datasets**

| Dataset      | Description  | References                |
|--------------|--|---------------------------|
| MM-COVID     | This dataset comprises 3981 fake news content and 7192 genuine news content collected from varied news sources of 6 different languages.   | (Y. Li et al., 2020)      |
| ReCOVery     | This dataset comprises textual, temporal, network, and visual information related to health misinformation tweets and articles collected from Reuters and public radio websites. | (Zhou et al., 2020)       |
| Twitter data | This dataset has 780 million tweets, from all over the world, in multiple languages.   | (Lopez & Gallemore, 2020) |
| CoAID        | Covid-19 Healthcare Misinformation Dataset (CoAID), has fake news tweets along with users' social engagement about such news.  | (Cui & Lee, 2020)         |

**Future Research Direction and Questions**

**Table 3. Future Research Questions according to identified concepts.**

| Concept areas and Research questions  |
|---|
| <p>Feature-Oriented</p> <ul style="list-style-type: none"> <li>How do User-centric features affect the tendency to share health misinformation?</li> <li>How do Linguistic centric features affect the tendency to share health misinformation?</li> <li>How do user engagement features affect the tendency to share health misinformation?</li> <li>How do Lexicon-based features affect the tendency to share health misinformation?</li> <li>What are the visual features to be considered to detect health misinformation on visual social media platforms?</li> <li>What are the writing style features to be considered to detect health misinformation on visual social media platforms?</li> <li>What are the Audio features to be considered to detect health misinformation on visual social media platforms?</li> </ul> |
| <p>Technique-Oriented</p> <ul style="list-style-type: none"> <li>Are there any reinforcement learning techniques present to detect health misinformation?</li> <li>Are there any transfer learning techniques present to detect health misinformation?</li> <li>Are there any generative adversarial neural network techniques present to detect health misinformation?</li> </ul>  |
| <p>Generalized</p> <ul style="list-style-type: none"> <li>What features help to eliminate health misinformation from social media?</li> <li>What are the techniques present to eliminate misinformation from social media?</li> <li>How to perform credibility assessment of health information on social media platforms?</li> </ul>   |

Based on the existing literature review, we noticed that the detection and elimination of proliferated health misinformation is an evolving area and has great scope to explore further. Therefore, we categorized the identified research gaps among three broad themes as Feature-based enhancement, Technique-based enhancement, and generalized. Proliferated health misinformation on social media is often present in textual forms (posts, tweets, blogs). Hence many researchers such as (Glasdam & Stjernswärd, 2020; Porat et al., 2019; Sear et al., 2020; Silva et al., 2020) have focused on textual features (ex., n-grams, part of speech, etc.) whereas writing style, audio, and visual features also play an important role in the dissemination of misinformation. Furthermore, researchers such as (Cui et al., 2020; Cui & Lee, 2020; Ghenai & Mejova, 2018; Y. Li et al., 2020) utilized a wide variety of techniques such as qualitative/quantitative surveys, machine learning algorithms, topic modeling, analytical tools, and statistical methods. Whereas other state-

of-the-art deep learnings, transfer learning, and reinforcement learning techniques have shown promising results in the area of healthcare-associated problems such as medical imaging and noncredible lab reporting. In the future, researchers have a wide scope to adopt this problem, they can objectify their research as a quantitative, qualitative, and design science by utilizing the aforementioned gaps. Researchers can embrace any feature set and study their effects on user's willingness to spread health misinformation. Moreover, they can adopt any technique as a tool from the wide variety as presented in the literature review. Hence, we highlighted a series of research questions centered on three broad research themes – the detection techniques, features, and generalized research problems. See Table 3.

As shown in Table 3, the researchers may wish to address the research questions such as “Are there any transfer learning techniques present to detect health misinformation?” and “What are the visual features to be considered to detect health misinformation on visual social media platforms?”. To achieve the objective of automatically detecting the visual misinformation on Twitter, a dataset presented by (Zhou et al., 2020) can be used. The availability of the dataset makes it feasible to perform the research in this direction.

Theoretically, the proposed conceptual framework fosters the quest of detecting health misinformation by providing a variety of aspects to guide future research. Our systematic review provides a comprehensive understanding of the features and techniques adopted by prior studies to identify and eliminate health misinformation. Furthermore, we have listed the 13 future research questions that are categorized as feature-oriented, technique-oriented, and general inquiry. Practically, this systematic review is a guide to various stakeholders in the healthcare domain to consider studied features and techniques to effectively understand and detect the proliferated health misinformation that in part refrains the dissemination of false knowledge regarding health issues such as cure to treatable/untreatable diseases.

## Conclusion

The massive amount of available healthcare information on social media makes it difficult to distinguish reliable facts from fiction and the proliferation of health misinformation possesses many forms. Researchers are adopting various approaches to addressing this problem that requires more multidimensional approaches. In order to apply distinct techniques to detect health misinformation, an enormous amount of data is required to train the machine learning and deep learning-based models. Hence, we have found four distinct labeled datasets that will help future researchers in the pursuit of health misinformation elimination. The identified gaps and future research questions will help to broaden the scope of future health misinformation detection research works.

## References

- Abul-Fottouh, D., Song, M. Y., & Gruz, A. (2020). Examining algorithmic biases in YouTube's recommendations of vaccine videos. *International Journal of Medical Informatics*, 140, 104175. <https://doi.org/10.1016/j.ijmedinf.2020.104175>
- Barua, Z., Barua, S., Aktar, S., Kabir, N., & Li, M. (2020). Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. *Progress in Disaster Science*, 8, 100119. <https://doi.org/10.1016/j.pdisas.2020.100119>
- Boudewyns, V., Betts, K. R., Johnson, M., Paquin, R. S., O'Donoghue, A. C., & Southwell, B. G. (2020). Experimental evidence of consumer and physician detection and rejection of misleading

- prescription drug website content. *Research in Social and Administrative Pharmacy*. <https://doi.org/10.1016/j.sapharm.2020.06.019>
- Christozov, D., & Mitreva, E. (2020). TRUST IN LEARNING FROM BIG DATA: THE TWO SIDES OF THE SAME COIN. *Issues in Information Systems*, 21(1).
- Cui, L., & Lee, D. (2020). CoAID: COVID-19 Healthcare Misinformation Dataset. *ArXiv:2006.00885 [Cs]*. <http://arxiv.org/abs/2006.00885>
- Cui, L., Seo, H., Tabar, M., Ma, F., Wang, S., & Lee, D. (2020). DETERRENT: Knowledge Guided Graph Attention Network for Detecting Healthcare Misinformation. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 492–502. <https://doi.org/10.1145/3394486.3403092>
- Datta, R., Yadav, A. K., Singh, A., Datta, K., & Bansal, A. (2020). The infodemics of COVID-19 amongst healthcare professionals in India. *Medical Journal Armed Forces India*, 76(3), 276–283. <https://doi.org/10.1016/j.mjafi.2020.05.009>
- Ghenai, A. (2017). Health Misinformation in Search and Social Media. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1371–1371. <https://doi.org/10.1145/3077136.3084153>
- Ghenai, A., & Mejova, Y. (2018). Fake Cures: User-centric Modeling of Health Misinformation in Social Media. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), 1–20. <https://doi.org/10.1145/3274327>
- Glasdam, S., & Stjernswärd, S. (2020). Information about the COVID-19 pandemic – A thematic analysis of different ways of perceiving true and untrue information. *Social Sciences & Humanities Open*, 2(1), 100090. <https://doi.org/10.1016/j.ssaho.2020.100090>
- Hou, R., Perez-Rosas, V., Loeb, S., & Mihalcea, R. (2019). Towards Automatic Detection of Misinformation in Online Medical Videos. *2019 International Conference on Multimodal Interaction*, 235–243. <https://doi.org/10.1145/3340555.3353763>
- Islam, A. K. M. N., Laato, S., Talukder, S., & Sutinen, E. (2020). Misinformation sharing and social media fatigue during COVID-19: An affordance and cognitive load perspective. *Technological Forecasting and Social Change*, 159, 120201. <https://doi.org/10.1016/j.techfore.2020.120201>
- Kawchuk, G., Hartvigsen, J., Innes, S., Simpson, J. K., & Gushaty, B. (2020). The use of internet analytics by a Canadian provincial chiropractic regulator to monitor, evaluate and remediate misleading claims regarding specific health conditions, pregnancy, and COVID-19. *Chiropractic & Manual Therapies*, 28(1), 24. <https://doi.org/10.1186/s12998-020-00314-9>
- Koohang, A., Paliszkievicz, J., & Nord, J. H. (2018). SOCIAL MEDIA PRIVACY CONCERNS AMONG COLLEGE STUDENTS. *Issues in Information Systems*, 19(1).
- Kouzy, R., Abi Jaoude, J., Kraitem, A., El Alam, M. B., Karam, B., Adib, E., Zarka, J., Traboulsi, C., Akl, E., & Baddour, K. (2020). Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter. *Cureus*. <https://doi.org/10.7759/cureus.7255>

- Lavorgna, L., Stefano, M. D., Sparaco, M., Moccia, M., Abbadessa, G., Montella, P., Buonanno, D., Esposito, S., Clerico, M., Cenci, C., Trojsi, F., Lanzillo, R., Rosa, L., Morra, V. B., Ippolito, D., Maniscalco, G., Bisecco, A., Tedeschi, G., & Bonavita, S. (2018). Fake news, influencers and health-related professional participation on the Web: A pilot study on a social-network of people with Multiple Sclerosis. *Multiple Sclerosis and Related Disorders*, 25, 175–178. <https://doi.org/10.1016/j.msard.2018.07.046>
- Li, J., Chen, W.-H., Xu, Q., Shah, N., Kohler, J. C., & Mackey, T. K. (2020). Detection of self-reported experiences with corruption on twitter using unsupervised machine learning. *Social Sciences & Humanities Open*, 2(1), 100060. <https://doi.org/10.1016/j.ssaho.2020.100060>
- Li, Y., Jiang, B., Shu, K., & Liu, H. (2020). *MM-COVID: A Multilingual and Multidimensional Data Repository for Combating COVID-19 Fake New*.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Journal of Clinical Epidemiology*, 62(10), e1–e34.
- Lopez, C. E., & Gallemore, C. (2020). *An Augmented Multilingual Twitter Dataset for Studying the COVID-19 Infodemic* [Preprint]. In Review. <https://doi.org/10.21203/rs.3.rs-95721/v1>
- Mądra-Sawicka, M., Paliszkievicz, J., & Nord, J. H. (2020). Determinants of Social Media Usage in Business by Women: Age and Development of the Country. *Information*, 11(9), 445.
- Niknam, F., Samadbeik, M., Fatehi, F., Shirdel, M., Rezazadeh, M., & Bastani, P. (2020). COVID-19 on Instagram: A content analysis of selected accounts. *Health Policy and Technology*. <https://doi.org/10.1016/j.hlpt.2020.10.016>
- Oberiri, D. A., & Omar, B. (2020). User motivation in fake news sharing during the COVID-19 pandemic: An application of the uses and gratification theory. *Online Information Review*, 45(1), 220–239.
- Porat, T., Garaizar, P., Ferrero, M., Jones, H., Ashworth, M., & Vadillo, M. A. (2019). Content and source analysis of popular tweets following a recent case of diphtheria in Spain. *European Journal of Public Health*, 29(1), 117–122.
- Sear, R. F., Velasquez, N., Leahy, R., Restrepo, N. J., Oud, S. E., Gabriel, N., Lupu, Y., & Johnson, N. F. (2020). Quantifying COVID-19 Content in the Online Health Opinion War Using Machine Learning. *IEEE Access*, 8, 91886–91893. <https://doi.org/10.1109/ACCESS.2020.2993967>
- Secosan, I., Virga, D., Crainiceanu, Z. P., Bratu, L. M., & Bratu, T. (2020). Infodemia: Another Enemy for Romanian Frontline Healthcare Workers to Fight during the COVID-19 Outbreak. *Medicina*, 56(12), 679. <https://doi.org/10.3390/medicina56120679>
- Shi, S., Brant, A. R., Sabolch, A., & Pollom, E. (2019). False News of a Cannabis Cancer Cure. *Cureus*. <https://doi.org/10.7759/cureus.3918>

- Silva, M., Ceschin, F., Shrestha, P., Brant, C., Fernandes, J., Silva, C. S., Grégio, A., Oliveira, D., & Giovanini, L. (2020). Predicting Misinformation and Engagement in COVID-19 Twitter Discourse in the First Months of the Outbreak. *ArXiv:2012.02164 [Cs]*. <http://arxiv.org/abs/2012.02164>
- Smaldone, F., Ippolito, A., & Ruberto, M. (2020). The shadows know me: Exploring the dark side of social media in the healthcare field. *European Management Journal*, 38(1), 19–32. <https://doi.org/10.1016/j.emj.2019.12.001>
- Tasnim, S., Hossain, M. M., & Mazumder, H. (2020). Impact of Rumors and Misinformation on COVID-19 in SocialMedia. *Journal of Preventive Medicine and Public Health*, 53(3), 171–174. <https://doi.org/10.3961/jpmp.20.094>
- Troshani, I., & Wickramasinghe, N. (2018). Contemporary Developments in e-Health. In *Theories to Inform Superior Health Informatics Research and Practice* (pp. 391–401). Springer.
- Venegas-Vera, A. V., Colbert, G. B., & Lerma, E. V. (2020). Positive and negative impact of social media in the COVID-19 era. *Reviews in Cardiovascular Medicine*, 21(4), 561–564. <https://doi.org/10.31083/j.rcm.2020.04.195>
- Vyas, P., & El-Gayar, O. (2020). Credibility Analysis of News on Twitter using LSTM: An exploratory study. *AMCIS 2020 Proceedings*, 17.
- Waszak, P. M., Kasprzycka-Waszak, W., & Kubanek, A. (2018). The spread of medical fake news in social media – The pilot quantitative study. *Health Policy and Technology*, 7(2), 115–118. <https://doi.org/10.1016/j.hlpt.2018.03.002>
- Wong, L.-W., Tan, G. W.-H., Hew, J.-J., Ooi, K.-B., & Leong, L.-Y. (2020). Mobile social media marketing: A new marketing channel among digital natives in higher education? *Journal of Marketing for Higher Education*, 1–25. <https://doi.org/10.1080/08841241.2020.1834486>
- Zhou, X., Mulay, A., Ferrara, E., & Zafarani, R. (2020). ReCOVerry: A Multimodal Repository for COVID-19 News Credibility Research. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 3205–3212. <https://doi.org/10.1145/3340531.3412880>