

Access this article online

Quick Response Code:



Website:
www.jehp.net

DOI:
10.4103/jehp.jehp_1136_20

A content analysis of social media on outbreak of infodemic amidst COVID-19 pandemic

Anusha Divvi, Shivashankar Kengadaran, Lakshmi Subhashini Katuri¹, M. Kavitha², Arunkumar Sundaragopal³, Vani Anusha⁴, Sivabalakumaran Kengadaran⁵, Hemachandran Sekar⁶

Department of Public Health Dentistry, Indira Gandhi Institute of Dental Sciences, Sri Balaji Vidyapeeth, Puducherry, India, ¹Department of Medicine, Geetam Institute of Medical Sciences and Research, Visakhapatnam, Andhra Pradesh, India, ²Department of Paediatric and Preventive Dentistry, Indira Gandhi Institute of Dental Sciences, Sri Balaji Vidyapeeth Puducherry, India, ³Department of Public Health Dentistry, Sri Venkateshwara Dental College and Hospitals, Chennai, Tamil Nadu, India, ⁴Department of Oral Medicine and Radiology, Indira Gandhi Institute of Dental Sciences, Sri Balaji Vidyapeeth Puducherry, India, ⁵Department of Anesthesia, Chengalpattu Government Medical College and Hospital, Tamil Nadu, India, ⁶Department of Medicine, Indira Gandhi Medical College and Research Institute, Puducherry, India

Address for correspondence:

Dr. Shivashankar Kengadaran,
Department of Public Health Dentistry,
Indira Gandhi Institute of Dental Sciences,
Puducherry - 607 402,
India.
E-mail: shiva.freee@gmail.com

Received: 01-09-2020
Accepted: 24-12-2020
Published: 29-10-2021

Abstract:

BACKGROUND: With so much content on social media platforms about COVID-19, determining which information is reliable can be a daunting task. Hence, this study is aimed to analyze various posts with regard to COVID-19 on various social media platforms for their reliability and also examined various factors that influence information reliability.

MATERIALS AND METHODS: A cross-sectional study was conducted, with 934 samples related to coronavirus pandemic published on Twitter, Instagram, and Facebook using systematic random sampling. We adopted the criteria given by Paul Bradshaw and modified to assess the characteristics of the samples. Training and calibration of the investigators were carried out for 3 consecutive days before beginning the study. The data were analyzed using the Chi-square test and multinomial logistic regression to estimate the odds ratios.

RESULTS: Out of 934 samples studied, only 570 (61%) were found to be reliable of which 243 (42.6%) were from Twitter, 117 (20.6%) from Instagram, and 210 (36.8%) from Facebook. We found that the reliability of the information on social media platforms is significantly influenced by network (odds: 1.32; 95% confidence interval [CI]: 1.16–1.52; $P = 0.036$), content (odds: 1.83; 95% CI: 1.69–1.92; $P = 0.009$), contextual update (odds: 1.41; 95% CI: 1.24–1.53) and age of the account (odds: 1.92; 95% CI: 1.64–2.09; $P = 0.002$).

CONCLUSION: Our study shows that the reliability of the social media posts significantly depends on the network, contextual update, and age of the account. Hence, cross verifying the information from a reliable source is the need of the hour to prevent panic and mental distress.

Keywords:

Coronavirus, information, misinformation, social media

Introduction

The outbreak of coronavirus disease is declared as a public health emergency of worldwide concern by the World Health Organization (WHO). As of April 7, 2020, 13,48,628 cases of coronavirus disease 2019 (COVID-19), including 74,816 deaths, had been reported globally.^[1] As the world combats this pandemic, a global epidemic of misinformation is spreading rapidly through various social media

platforms, which in turn poses a serious threat to the mental health of the public. The WHO Director-General Tedros Adhanom Ghebreyesus said “We’re not just fighting an epidemic; we’re fighting an infodemic,” at the Munich security conference.^[2]

Every outbreak will be flooded with abundant information, which also includes rumors and fake information. This was evident even within the middle ages. However, the difference currently with social networking is that this phenomenon

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Divvi A, Kengadaran S, Katuri LS, Kavitha M, Sundaragopal A, Anusha V, *et al.* A content analysis of social media on outbreak of infodemic amidst COVID-19 pandemic. *J Edu Health Promot* 2021;10:366.

is intensified, it spreads faster and further, like viruses.^[3] This infodemic compromises outbreak response and increase public confusion; generates fear and panic due to unsubstantiated rumors and overstated claims.^[4]

In this digital era, the time needed to analyze, assess, and communicate information cannot compete with the instantaneous spreading of misinformation on social media platforms. False information spread easily through the web or any social media platforms and impact not only the mental health of the users but also engenders economy throughout the world and by the time it is identified as false information, it creates irreparable chaos.^[5] For instance, a false rumor that the coronavirus can be spread by chickens has slashed Indian poultry sales by almost 50%.^[6] Various government agencies are now setting up services to debunk stories that they consider to be false and not credible.

Information on social media platforms is a double-ended sword, suffers from a relative lack of professional gatekeepers to monitor the content. Among the lockdown due to the sudden outbreak of coronavirus disease, people have turned to social media to pass their time.^[7] To evaluate the information credibility on social media has become an important issue today. Hence, this study aims to analyze various posts about COVID-19 on various social media platforms for their reliability and also examined various factors that influence information credibility.

Materials and Methods

Study design and setting: Cross-sectional study and Social Media Platforms

Study participants and sampling: A study across 45-min composite time period was conducted, with the posts and tweets related to coronavirus pandemic published on social media platforms between March 1, 2020, and March 20, 2020. Three social media platforms (Facebook, Instagram, and Twitter) were randomly selected from the list of most commonly used social media platforms.^[8] At three randomly generated time slots, the search was done using the key terms #COVID2019, #Coronavirus, #COVID19, and #Corona across Facebook, Instagram, and Twitter. All the posts were sorted from the latest date and extracted using screenshot methods over the next 15 min. The posts with completely irrelevant content were excluded from the study.

Data collection tool and technique

Coding scheme

The study adopted criteria given by Bradshaw^[9] and modified to assess the characteristics of the samples

concerning the following variables: (i) Location of the source or account, (ii) Network of the account, (iii) Content of the post or tweet, (iv) Contextual updates, (v) Age of the account (<6 months and >6 months), and (vi) Reliability of the information and source. Each of the variables was coded as either 1 (positive) or 0 (negative) except for the age of the account which was a continuous variable.

Coding procedure

Two investigators were involved in assessing and coding the posts and tweets. These investigators were trained and calibrated for 3 consecutive days before the start of the study. Training sessions involved several group discussions about the meanings and nuances of coding categories. Each investigator was asked to assess and code ten samples from each social media platform. Inter- and Intra-investigator reliability was assessed using Kappa statistics. The average Inter- and Intra-investigator reliability scores were 0.86 and 0.94, respectively [Table 1].

Ethical consideration: The study was approved by Institutional ethical committee, IGIDS

Statistical analysis

The statistical analyses were performed using the Statistical Package for the Social Science (SPSS, Inc., Chicago, IL; USA, version, 21.00). Chi-square analysis was done to assess the differences in the reliability of samples and various factors influencing it. A binomial logistic regression analysis was fitted to estimate the odds ratios (ORs). The 95% confidence interval (CI) was used to examine the possible factors influencing the reliability of information published on social networking websites and the results were obtained by comparing them with the reference category. The $P < 0.05$ was considered to be statistically significant.

Results

Across the 45-min composite study period, 986 posts were retrieved from the three social media platforms of which 52 were completely irrelevant. The remaining 934 samples were analyzed, of which 443 (47.4%) samples were those posted on Twitter, 187 (20%) on Instagram, and 304 (32.6%) on Facebook. The distribution of samples in the three social networking websites concerning various characteristics including location, network, content, contextual update, age of the account, and reliability is mentioned in Table 2. The differences in the reliability of samples concerning the different variables such as location, network, content, contextual update, and age of the account are shown in Table 3. We observed that the reliability was influenced significantly by these variables ($P < 0.05$).

Table 1: Operational definition and Inter & Intra-examiner reliability of each variable

Variables	Facebook		Instagram		Twitter	
	Inter-examiner reliability	Intra-investigator reliability	Inter-investigator reliability	Intra-investigator reliability	Inter-investigator reliability	Intra-investigator reliability
Location of the source	0.82	0.94	0.84	0.96	0.94	0.92
Does the account holder belong to the same location of which they posts/tweets about?						
Network	0.78	0.90	0.86	0.96	0.86	0.90
Does the account have more than 100 followers?						
Is it a familiar or official account of a recognized organization or person?						
Content	0.92	0.92	0.88	0.94	0.82	0.94
Can the information be corroborated with other sources like WHO?						
Contextual update	0.86	0.94	0.86	0.96	0.9	0.88
Have the account holder posted similar health or social information before 2020?						
Age of the account	0.82	0.94	0.82	0.96	0.86	0.96
Number of months since the account has been created						
Reliability	0.86	0.96	0.86	0.92	0.86	0.92
Is the information in the account reliable or can be acted upon?						

WHO=World Health Organization

Table 2: Number of posts published on the three social network websites based on various factors

Factors	Group		
	Twitter, n (%)	Instagram, n (%)	Facebook, n (%)
Location			
Different	163 (36.8)	140 (74.9)	117 (38.5)
Same	280 (63.2)	47 (25.1)	187 (61.5)
Network			
≤ 100 followers	210 (47.4)	140 (74.9)	187 (61.5)
>100 followers	233 (52.6)	47 (25.1)	117 (38.5)
Content			
Cannot be corroborated	187 (42.2)	94 (50.2)	210 (69.1)
Can be corroborated	256 (57.8)	93 (49.8)	94 (30.9)
Contextual update			
No	233 (52.6)	164 (87.7)	234 (77.0)
Yes	210 (47.4)	23 (12.3)	70 (23.0)
Age of the account (months)			
≤6	163 (36.8)	70 (37.4)	94 (30.9)
>6	280 (63.2)	117 (62.6)	210 (69.1)
Reliability			
No	200 (36.80)	70 (37.4)	94 (30.9)
Yes	243 (63.2)	117 (62.6)	210 (69.1)

ORs with CIs were calculated for the data to observe how the variables including location, network, content, contextual update, and age influence the reliability of the information published on social networking websites.

The results indicated that the information on social networking accounts aged >6 months appeared to have higher significant odds (odds: 1.92; 95% CI: 1.64–2.09; $P = 0.002$) of being reliable [Table 4].

Table 3: Effect of various factors on the reliability of the posts published on the three social network websites

Factor	Number of reliable posts			P [#]
	Twitter, n (%)	Instagram, n (%)	Facebook, n (%)	
Location				
Different	106 (43.6)	48 (41)	117 (55.7)	0.036*
Same	137 (56.4)	69 (59)	138 (44.3)	
Network				
≤ 100 followers	42 (17.3)	19 (11.8)	13 (6.2)	>0.001*
>100 followers	201 (82.7)	98 (88.2)	197 (93.8)	
Content				
Cannot be corroborated	67 (27.6)	50 (42.7)	10 (4.8)	>0.001*
Can be corroborated	176 (72.4)	67 (57.3)	200 (95.2)	
Contextual update				
No	33 (13.6)	35 (30)	56 (26.7)	0.002*
Yes	210 (86.4)	82 (70)	154 (73.3)	
Age (months)				
≤ 6	54 (22.2)	25 (21.4)	31 (14.8)	0.001*
>6	189 (77.8)	92 (78.6)	179 (85.2)	

[#]Chi-square analysis, *P<0.05 is considered as statistically significant

Table 4: Factors influencing the reliability of information published on social networking websites

Factor	OR	95% CI	P
Location			
Different	1.00	Reference	0.068
Same	1.22	0.91-1.36	
Network			
≤ 100 followers	1.00	Reference	0.036*
>100 followers	1.32	1.16-1.52	
Content			
Cannot be corroborated	1.00	Reference	0.009*
Can be corroborated	1.83	1.69-1.92	
Contextual update			
No	1.00	Reference	0.012*
Yes	1.41	1.24-1.53	
Age (months)			
≤ 6	1.00	Reference	0.002*
>6	1.92	1.64-2.09	

*P<0.05 is considered as statistically significant. OR is calculated using binomial logistic regression analysis. Dependent variable: Reliability categorized as yes/no. OR=Odds' ratio, CI=Confidence interval

Apparently, the odds of information on social media platforms being reliable was significantly higher when the account holder has >100 followers (odds: 1.32; 95% CI: 1.16–1.52; $P = 0.036$). The information posted on the social communication platforms which can be corroborated with other sources showed significantly higher odds (odds: 1.83; 95% CI: 1.69–1.92; $P = 0.009$) in terms of reliability. When the effect of the contextual update was taken into consideration, the information posted by the social network account holders who had posted similar health or social information before 2020, exhibited higher odds (odds: 1.41; 95% CI: 1.24–1.53) of reliability. However, the location of the account holder did not found to have a significant influence on the reliability of the information [Table 4].

Discussion

The first case of COVID-19 was reported in China in December 2019. From then, along with the spread of the disease, there was an outbreak of infodemic creating chaos and panic among people. Further with lockdown or social isolation put into effect, there was an increase in social media traffic (13% in February to 16% in March) which was evident through the Social Media Index. Governments of various countries and health agencies like the WHO have issued several warnings regarding the spread of misinformation. However, it is not practically possible to track the person spreading the information amidst this crisis.

Twitter is popular with journalists, politicians, and celebrities, hence, many users turn to Twitter for trending news.^[10,11] This might be a possible explanation for a larger number of samples (47.4%) from this social communication platform compared to Facebook (32.6%) and Instagram (20%). One of the basic attributes of journalism is to cross-check information with reliable sources like legal documents and government or government-supported databases. It is considered that the information that matches with these sources is highly reliable.^[10] Interestingly, the results of the current study indicate that the odds of being reliable is higher for the information posted on the social communication platforms which can be corroborated with other sources.

According to Tong *et al.*, the number of friends or followers on a social media platform reflects the popularity of the profile's owner.^[12] In addition, according to Westerman *et al.*, it is likely that any individual follows another user on the social media platform mainly because their posts contain some reliable information.^[13] Thus, the number of followers of an account holder on a social communication

site is likely to positively influence the reliability of the information posted by him/her. Similarly, in our study, it was found that the information posted by an account holder with >100 followers is more reliable.

According to Bradshaw, checking the background history of the account holder, i.e., posts or tweets published by them probably helps to assess the credibility of the information posted by them.^[9] If the user had tweeted or posted about any health or social information previously it is more likely that the information they share currently is credible. Likewise, our findings also revealed contextual updates as a factor to influence the reliability of the information shared by the user on the social media platform.

Bradshaw also suggested that the newer the account is, the more skeptical we have to be about the information they share. It is more likely that new users may claim to have information on any current issue to gain the attention of the public.^[9] Similarly, we observed in our study that information on social networking accounts aged >6 months appeared to be significantly reliable.

Limitations and future research

The findings of our study should be understood in light of its limitations. The first limitation is a relatively small sample size due to which we could not capture any confounding effect in our study. Hence, we recommend future research with larger and more varied samples. Besides, we believe that our research represents empirical stages in the development of an extensive understanding of online information evaluation. Hence, one must be cautious in generalizing the findings of our study.

Conclusion

Based on the results, the reliability of the social media posts significantly depends on network, contextual update, and age of the account. While the growing usage of social media during social isolation and its potential to disseminate fake information, the ability to find reliable information is the need of the hour. Our findings provide an initial step in understanding what factors influence the reliability of the information on social media platforms so that the information providers and users can evaluate information reliability more effectively.

Summary points:

- All the health information available or posted in social media are not reliable

- The reader has to verify the health information before relying upon the post
- Information posted by healthcare-related personal and information from older accounts can be relied.

Acknowledgment

The authors would like to acknowledge Mr. Divvi Sriraj and Mrs. Sridevi Jampani for their technical support.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Kengadaran S., Divvi A., Kengadaran S. Prepare to succeed, panic to fail: The mantra for COVID-19 prevention in India. *Population Medicine*. 2020;2(April):7. doi: 10.18332/popmed/120074
2. WHO. WHO Director-General, Speeches, Munich Security Conference. Available from: <https://www.who.int/dg/speeches/detail/munich-security-conference>. [Last accessed on 2020 Apr 07].
3. Zarocostas J. How to fight an infodemic. *Lancet* 2020;395:676.
4. Brindha MD, Jayaseelan R, Kadeswara S. Social media reigned by information or misinformation about COVID-19: a phenomenological study. *SSRN Electronic Journal* April. 2020 Aug 23.
5. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. *Soc Sci Med* 2019;240:112552.
6. The Times of India. India Chicken Sales Slashed almost 50% by False Virus Rumor. Available from: <https://timesofindia.indiatimes.com/business/india-business/india-chicken-sales-slashed-almost-50-by-false-virus-rumour/articleshow/74372093.cms>. [Last accessed on 2020 Apr 07].
7. Li R, Suh A. Factors influencing information credibility on social media platforms: Evidence from Facebook pages. *Procedia Comput Sci* 2015;72:314-28.
8. Statista. Global Social Networks Ranked by Number of Users 2020. *Social Media & User-Generated Content*; 2020.
9. Bradshaw P. *The Online Journalism Handbook: Skills to Survive and Thrive in the Digital Age*. United Kingdom: Routledge; 2017.
10. Perloff RM. *The Dynamics of News: Journalism in the 21st-Century Media Milieu*. United Kingdom Routledge; 2019.
11. Lee SY. Effects of relational characteristics of an answerer on perceived credibility of informational posts on social networking sites: The case of Facebook. *Inf Res Int Electron J* 2018;23:n3.
12. Tong ST, Van Der Heide B, Langwell L, Walther JB. Too much of a good thing? The relationship between number of friends and interpersonal impressions on Facebook. *J Comput Mediat Commun* 2008;13:531-49.
13. Westerman D, Spence PR, Van Der Heide B. A social network as information: The effect of system generated reports of connectedness on credibility on Twitter. *Comput Hum Behav* 2012;28:199-206.