City, University of London Institutional Repository

Understanding the complex links between social media and health behaviour

Fabiana Zollo and colleagues call for comprehensive, robust research on the influence of social media on health behaviour in order to improve public health responses

Over 90% of people connected to the internet are active on social media, with a total of 4.76 billion users worldwide in January 2023.[1] The digital revolution has reshaped the news landscape and changed the way users interact with information. Social media's targeted communication rapidly reaches vast audiences, who in turn actively participate in shaping and engaging with content. This marks a departure from the more passive consumption patterns associated with traditional media.

Over the past few years, social media have emerged as a primary source of news for many people, despite widespread user concerns about potential misinformation (box 1) and the necessity to discern between reliable and untrustworthy information.[4] Data from six continents also indicate a preference among users for content that reflects their reading or viewing history, rather than content selected by journalists, suggesting a shift towards personalised and user driven content curation. In this evolving landscape, celebrities, influencers, and social media personalities are increasingly assuming roles as news sources, especially on platforms such as TikTok, Instagram, and Snapchat.

Public health organisations have recognised the crucial role of social

KEY MESSAGES

- Monitoring social media is important to understand public perceptions, biases, and false beliefs
- Drawing conclusions on how social media affects health behaviour is difficult because measures are unstandardised, sources are limited, and data are incomplete and biased
- Rigorous research is needed from varied settings and demographics to improve understanding of the effect of social media on health behaviour

Box 1: What is misinformation?

- The term "misinformation" is commonly used, yet its definitions can vary between studies, methods, and scholars, leading to disagreements on its precise meaning[2]
- Misinformation encompasses false, inaccurate, or misleading information, and is often distinguished from disinformation by its lack of deliberate creation and dissemination with the intent to deceive. Classifications may also extend to conspiracy theories and propaganda
- Misinformation poses a substantial risk to public health since it can undermine compliance with important public health measures such as vaccination uptake or physical distancing guidelines[3]

media in shaping the public debate and are working to utilise social media platforms to inform the public, combat misinformation, and improve health knowledge, attitudes, and behaviour. However, causal research on how social media information affects actual health behaviour is inconclusive, primarily because of methodological challenges associated with connecting online activity to offline actions and accurately measuring behavioural outcomes.[5] Thus, exploring the complex relations between information consumption, personal beliefs, and societal effects remains an important area of study. The development of vaccines against covid-19 was accompanied by an infodemic—an overabundance of information, not all of which is accurate.[6] Study of this phenomenon provides useful insight into the interplay between social media and health behaviours and the opportunities and challenges for research and practice.

Access to data on misinformation and health behaviour

Social media have provided unprecedented opportunities through which health information, including misinformation, can be amplified and spread. However, the impact of exposure to and interaction with misinformation on health behaviour remains a subject of debate within the scientific community.[7] While considerable evidence exists from research indicating that misinformation can affect knowledge, attitudes, or behavioural intentions, reaching a consensus in the scientific community on the links between social media and actual health behaviour has been challenging due to lack of data and inherent limitations in study design.

A recent systematic review of randomised controlled trials, for example, highlighted the need for more conceptual and theoretical work on the causal pathways through which misinformation shapes people's beliefs and behaviours.[8] This influence is often indirect, meaning that exposure to misinformation may affect changes in health behaviour by shaping psychological factors such as beliefs, feelings, and motivations (the so called psychological antecedents) which are commonly used to explain and predict behaviours. However, the roles of potential mediators such as emotions, social norms, and trust are still poorly understood. While all the studies in that review assessed the effect of misinformation on antecedents (intentions, attitudes, and subjective norms), only two of them measured actual behaviour. These studies included behavioural measures of activism, such as the act of signing petitions, yet none examined the effects of misinformation exposure on direct health measures or behaviours, such as vaccination. Indeed, the literature is unclear about the causal effect of individual online activity on behaviour. For example, while some research shows that risk perceptions and vaccination intentions can be affected by short visits to antivaccination websites,[9] exposure to antivaccination comments posted on news stories online appears to have little influence on individuals' perspectives regarding vaccines, although it could potentially undermine individuals'

1

trust in important health communication institutions.[10]

Furthermore, drawing links or establishing causality is not a trivial endeavour. One important obstacle to understanding the effect of social media on behaviour lies in the challenge of linking online activity with offline behaviour. This difficulty stems from factors such as data scarcity and privacy concerns, particularly regarding personal and sensitive information, that complicate efforts to assess how online interactions translate into real word actions. Establishing a clear connection between information consumption on social media platforms and tangible behavioural outcomes, while excluding the influence of external variables, is a complex task, especially when examining behaviour over medium to long term periods. Examining behaviour in the long term requires longitudinal data, which are often lacking due to the resources (such as costs and time) required for such research. Adding to these challenges is the lack of standardised measures and definitions across studies. As we have seen, misinformation is not unanimously defined (box 1), and health behaviours also encompass a variety of actions—in the covid-19 pandemic alone, behaviours ranged from adherence to nonpharmaceutical measures such as physical distancing to lockdowns, from handwashing to vaccinations. Even with clear definitions, measuring health behaviours reliably and accurately remains a considerable challenge.[11] Many studies rely heavily on self reported data, which may have a low correlation with objectively measured behaviour. Moreover, many differences exist between countries in terms of how data of this nature are collected. This variability makes it difficult to extrapolate definitive findings from different settings and contexts.

The covid-19 pandemic presented a unique opportunity to further investigate the potential effect of infodemics on health behaviour, especially on vaccine hesitancy and refusal. The evidence in the literature paints a complex picture of the relationship between social media misinformation and vaccination. On the one hand, researchers have identified a negative relationship between sharing misinformation online and vaccination uptake in the United States of America.[12] Similarly, a study in the UK and US suggests that exposure to misinformation reduces individuals' intention to vaccinate for their own and others' protection.[13]

These findings highlight the potentially detrimental effect of misinformation on public health efforts. A large review of 205 articles looked more specifically into conspiracies around vaccination (under review by journal not yet accepted). While some studies showed causal evidence of exposure to conspiracies on vaccination intentions, most studies were correlative and behaviour was not investigated. Thus, the findings of many studies suggest that uncertainty persists on the causality of this relationship. Further investigation into the association between social media behaviour and attitudes towards covid-19 vaccines showed that vaccine hesitancy was associated with interaction and consumption of low quality information online.[14] These results remained significant even after accounting for relevant variables, suggesting that social media behaviour may play an important role in predicting vaccine attitudes. Supporting this finding, a recent systematic review examining the role of social media as a predictor of covid-19 vaccine outcomes showed predominantly negative associations between social media predictors and vaccine perceptions, in particular concerning vaccine hesitancy. However, the evidence suggests a multifaceted landscape, with findings varying across different social media predictors, populations, and platforms.[15] Moreover, while concerns about infodemics shaping individuals' behavioural intentions are prevalent, some findings suggest a more nuanced reality. Despite the proliferation of information and debate on covid-19 vaccines, the relatively stable and positive trend in vaccine acceptance rates at an aggregated level challenges simplistic explanations about the effect of misinformation.[16]

Overall, despite the importance of the effect of social media and misinformation on health behaviour and the extensive assumptions within policy debates, the literature fails to provide definitive conclusions on a clear association between social media and health behaviour. As discussed earlier, measuring behavioural change is challenging due to the scarcity of studies incorporating actual behavioural measures, limitations in laboratory experiments, and difficulties in establishing connections with online activity. Studies are often confined to specific geographical areas, primarily western countries (notably the US), or limited to specific time periods. In addition, data samples are often constrained by the lack of comprehensive information, such as demographics or

geolocation. Furthermore, longitudinal studies are required with extensive access to social media behaviour as well as access to actual behavioural data.

Therefore, further studies are needed to assess the causal effect of social media on offline behaviour. This will require overcoming the ethical issues of data linkage and protection. Such studies will need to integrate social media data with information from different sources, adjusting statistical methods to handle sampling biases, and accounting for the inherent dynamics of social media discussions, which are often characterised by extreme polarisation and user segregation.

Social media dynamics and health
Social media debates are often marked by intense segregation. Users tend to seek out information that aligns with their existing beliefs while dismissing opposing viewpoints. Social media platforms, especially those employing content filtering algorithms, tend to exploit this natural tendency by favouring content aligned to the user's history and preferences.[17] After all, platforms such as Facebook are built on the foundational unit of the "like," which represents the most fundamental action a user can take within the environment. Selective exposure to like minded content can contribute to the formation of echo chambers—that is, well separated groups of like minded users—where individuals are surrounded by others who share similar opinions. This phenomenon can act as a breeding ground for the spread of misinformation and hinder its correction.[18] Analysis of Facebook users has shown the existence of opposing and separate communities—provaccine and antivaccine—with the latter group generally being more active.[19] Another study on the public discussion on covid-19 vaccines found similar results, showing users' inclination to interact with like minded individuals and the presence of segregated communities, with antivaccine groups exhibiting greater cohesiveness and stability over time.[20] Recent research has found that, on Facebook, like minded sources—that is, sources that align with users' political leanings—are indeed prevalent in what people see on the platform, although they do not seem to affect polarisation. In other words, no measurable effect on polarisation was seen when exposure to content from like minded sources was reduced.[21] Echo chambers and user segregation are crucial factors, as provaccination campaigns, for example, may become confined to individu-

als who already support vaccination, thus limiting their overall effectiveness. Recent research has explored how users engaged with covid-19 information on social media, and how such engagement changed over time.[22] Despite earlier findings suggesting that false news might spread faster than trustworthy information, analysis of various platforms indicates no substantial difference in the spread of reliable versus questionable information. Posts and interactions with misinformation sources follow similar growth patterns to those of reliable ones, although scaling factors specific to the platform apply. Mainstream platforms and Reddit have a smaller proportion of posts from questionable sources relative to reliable ones, while Gab stands out by notably amplifying posts from questionable sources. These results suggest that the primary drivers behind the spreading of reliable information and misinformation are the specific rules of the platform and the behaviours of groups and individuals engaged in the discussion, rather than the nature of the content.

Opportunities and challenges for using social media to improve health

The public actively engages in public debates through social media platforms based on their prior perceptions and beliefs. Identity is important; the extent to which people identify with their vaccination status is linked to the way the social media platforms are used. People who identify more strongly with being unvaccinated are less likely to use traditional news sources and rely more on information from social media and messaging services.[23] In this context, monitoring social media has become an essential and powerful tool for a dynamic and real time understanding of the information available to large parts of the public, their perceptions, and the presence of biases and false beliefs. The vast amount of data generated online enables the exploration and analysis of sociocognitive factors underlying the consumption and processing of information. When examined and aggregated, these data can provide valuable insights and reveal hidden patterns on people's perspectives. These insights can, in turn, support public communication efforts, ranging from monitoring public sentiment, concerns, and reactions to helping identify the informational needs of the population. Ultimately, this information can drive the development of recommendations aimed at improving the effectiveness of communication strategies and health measures. For

instance, a recent World Health Organization manual offers a guide to addressing the gap between health guidance recommendations and population behaviour using social listening.[24] Social media sources can be used to respond to specific question of concern, such as understanding why a certain community remains undervaccinated despite widespread availability of vaccines and strong recommendations for vaccination. This approach may facilitate a deeper understanding of the information environment of the population, their behaviour in seeking health information, and their health behaviours, thereby enabling the development of tailored strategies and recommendations.

Social media analyses usually rely on large amounts of data. However, it is important to acknowledge that these data may relate to unrepresentative segments of the population.[25] Therefore, it is crucial to pay careful attention to sample creation, which involves selecting a smaller subset of data from a larger population using a predefined selection method. This statistical challenge, known as sample selection bias, must be duly considered when seeking information about the overall population or specific groups who are less inclined to use social media. Although often oversimplified, social media presents a varied landscape, and the extent of sample selection problems may vary across countries and platforms.[26] For instance, Facebook usually covers a broader spectrum of the population in terms of both audience size and diversity of social groups, while X (formerly Twitter) and TikTok predominantly cater to specific subgroups, such as professionals and younger individuals. Additionally, the varying levels of user engagement in actively participating in conversations on social networks through comments, posts, likes, and other forms of interaction can also contribute to sample bias. Combining social data with information from other sources (for example, census data, electoral rolls, surveys, and health data) and employing statistical methods to adjust for sampling biases is thus crucial to obtain solid research outcomes (see, in another context, previous work on the Brexit referendum[27]). Such considerations are important for health communication campaigns that are inclusive and resonate with target audiences.

Learning the lessons

The covid-19 pandemic has heightened concerns about the potential effect of

misinformation in posing risks to public health. Yet, the issue extends beyond the recent crisis, and is important in shaping our response to future pandemics. Ensuring dissemination of accurate information is essential not only to safeguard public health in the present but also to mitigate risks and enhance preparedness for potential future crises. Assessing the effect of social media use on health behaviour is a complex task, with current evidence yet to be consolidated. To avoid biased outcomes, a comprehensive, multidimensional, and causal approach is necessary when investigating the interplay between online information and real world behaviour. Understanding causal relationships and their drivers will allow interventions to be developed to reduce the detrimental effects of online information on health. It is also essential to define clear outcomes. Indeed, online information about health can cover various aspects, including the formation of public opinion, effects on public discourse and agenda setting, interactions between doctors and patients, as well as influences on health behaviours in the short, medium, or long term.[28]

Further research is required to identify vulnerable populations and gain a better understanding of sociodemographic and ideological factors influencing users' behaviour. Additionally, cultural differences in information consumption and behaviours must also be considered to develop targeted and effective interventions and mitigate the influence of health misinformation.[29]

Addressing these questions requires robust data and study designs, with collaboration from digital platforms being crucial in accessing such data. A recent cooperation with Meta[30] allowed researchers to conduct multiple experiments and provided extensive access to user data from Facebook and Instagram. However, the success of this model relies entirely on the willingness of social media companies to participate. This highlights the need for an ethical, transparent collaboration and advocates for the democratisation of social media research through equitable data access.[31] Future studies should replicate these efforts in contexts other than politics, such as health, and expand research beyond the US to achieve a more comprehensive understanding of the effect of social media on behaviour globally.

Fabiana Zollo, professor[1,2]

Andrea Baronchelli, professor[3]

Cornelia Betsch, professor[4,5]

Marco Delmastro, research fellow[1,6]

Walter Quattrociocchi, professor[7]

[1]Ca' Foscari University of Venice, Venice, Italy

[2]New Institute Centre for Environmental Humanities, Venice, Italy

[3]City University of London, London, United Kingdom

[4]University of Erfurt, Erfurt, Germany

[5]Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany

[6]Enrico Fermi Center for Study and Research, Rome, Italy

[7]Sapienza University, Rome, Italy

Correspondence to: F Zollo
fabiana.zollo@unive.it

Check for updates

1. Kemp S. Digital 2023: global overview report. DataReportal—Global Digital Insights, 2023. https://datareportal.com/reports/digital-2023-global-overview-report.
2. Kozyreva A, Lewandowsky S, Hertwig R. Citizens versus the internet: confronting digital challenges with cognitive tools. *Psychol Sci Public Interest* 2020;21:103-56. doi:10.1177/1529100620946707
3. van der Linden S. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nat Med* 2022;28:460-7. doi:10.1038/s41591-022-01713-6
4. Reuters Institute digital news report. 2023. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf.
5. Budd J, Miller BS, Manning EM, et al. Digital technologies in the public-health response to COVID-19. *Nat Med* 2020;26:1183-92. doi:10.1038/s41591-020-1011-4
6. World Health Organization. Infodemic. 2023. https://www.who.int/health-topics/infodemic.
7. Briand SC, Cinelli M, Nguyen T, et al. Infodemics: a new challenge for public health. *Cell* 2021;184:6010-4. doi:10.1016/j.cell.2021.10.031
8. Schmid P, Altay S, Scherer LD. The psychological impacts and message features of health misinformation. *Eur Psychol* 2023;28:162-72. doi:10.1027/1016-9040/a000494
9. Betsch C, Renkewitz F, Betsch T, Ulshöfer C. The influence of vaccine-critical websites on perceiving vaccination risks. *J Health Psychol* 2010;15:446-55. doi:10.1177/1359105309353647
10. Dixon G. Undermining credibility: the limited influence of online comments to vaccine-related news stories. *J Health Commun* 2020;25:943-50. doi:10.1080/10810730.2020.1865485
11. Conner M, Norman P. Health behaviour: current issues and challenges. *Psychol Health* 2017;32:895-906. doi:10.1080/08870446.2017.1336240
12. Pierri F, Perry BL, DeVerna MR, et al. Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. *Sci Rep* 2022;12:5966. doi:10.1038/s41598-022-10070-w
13. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nat Hum Behav* 2021;5:337-48. doi:10.1038/s41562-021-01056-1
14. Rathje S, He JK, Roozenbeek J, Van Bavel JJ, van der Linden S. Social media behavior is associated with vaccine hesitancy. *PNAS Nexus* 2022;1:pgac207. doi:10.1093/pnasnexus/pgac207
15. McKinley CJ, Limbu Y. Promoter or barrier? Assessing how social media predicts Covid-19 vaccine acceptance and hesitancy: a systematic review of primary series and booster vaccine investigations. *Soc Sci Med* 2024;340:116378. doi:10.1016/j.socscimed.2023.116378
16. Valensise CM, Cinelli M, Nadini M, Galeazzi A, Peruzzi A, Etta G, et al. Lack of evidence for correlation between COVID-19 infodemic and vaccine acceptance. arXiv 2021:2107.07946. [Preprint.] doi:10.48550/arXiv.2107.07946
17. Cinelli M, De Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M. The echo chamber effect on social media. *Proc Natl Acad Sci U S A* 2021;118:e2023301118. doi:10.1073/pnas.2023301118
18. Zollo F, Bessi A, Del Vicario M, et al. Debunking in a world of tribes. *PLoS One* 2017;12:e0181821. doi:10.1371/journal.pone.0181821
19. Schmidt AL, Zollo F, Scala A, Betsch C, Quattrociocchi W. Polarization of the vaccination debate on Facebook. *Vaccine* 2018;36:3606-12. doi:10.1016/j.vaccine.2018.05.040
20. Santoro A, Galeazzi A, Scantamburlo T, Baronchelli A, Quattrociocchi W, Zollo F. Analyzing the changing landscape of the Covid-19 vaccine debate on Twitter. *Soc Netw Anal Min* 2023;13:115. doi:10.1007/s13278-023-01127-3
21. Nyhan B, Settle J, Thorson E, et al. Like-minded sources on Facebook are prevalent but not polarizing. *Nature* 2023;620:137-44. doi:10.1038/s41586-023-06297-w
22. Cinelli M, Quattrociocchi W, Galeazzi A, et al. The COVID-19 social media infodemic. *Sci Rep* 2020;10:16598. doi:10.1038/s41598-020-73510-5
23. Henkel L, Sprengholz P, Korn L, Betsch C, Böhm R. The association between vaccination status identification and societal polarization. *Nat Hum Behav* 2023;7:231-9. doi:10.1038/s41562-022-01469-6
24. World Health Organization, United Nations Children's Fund. How to build an infodemic insights report in six steps. 2023. https://www.who.int/publications/i/item/9789240075658.
25. Olteanu A, Castillo C, Diaz F, Kıcıman E. Social data: biases, methodological pitfalls, and ethical boundaries. *Front Big Data* 2019;2:13. doi:10.3389/fdata.2019.00013
26. Boyd D, Crawford K. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Inf Commun Soc* 2012;15:662-79. doi:10.1080/1369118X.2012.678878
27. Grčar M, Cherepnalkoski D, Mozetič I, Kralj Novak P. Stance and influence of Twitter users regarding the Brexit referendum. *Comput Soc Netw* 2017;4:6. doi:10.1186/s40649-017-0042-6
28. 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. doi:10.1016/j.socscimed.2019.112552
29. Kozyreva A, Lorenz-Spreen P, Herzog S, Ecker U, Lewandowsky S, Hertwig R. Toolbox of interventions against online misinformation and manipulation. *PsyArXiv* 2022. [Preprint.] doi:org/10.31234/osf.io/x8ejt
30. Tollefson J. Tweaking Facebook feeds is no easy fix for polarization, studies find. *Nature* 2023 Jul 27. doi:10.1038/d41586-023-02420-z
31. Roozenbeek J, Zollo F. Democratize social-media research—with access and funding. *Nature* 2022;612:404-404. doi:10.1038/d41586-022-04407-8

Cite this as: *BMJ* 2024;385:e075645
http://dx.doi.org/10.1136/bmj-2023-076478