

### **Perspective**

# Social media and the spread of misinformation: infectious and a threat to public health

Emily Denniss<sup>1,2,3,\*</sup>, and Rebecca Lindberg<sup>1,3</sup>

<sup>1</sup>School of Exercise and Nutrition Sciences, Deakin University, 221 Burwood Hwy, Burwood, Victoria 3125, Australia <sup>2</sup>School of Health and Social Development, Deakin University, 221 Burwood Hwy, Burwood, Victoria 3125, Australia

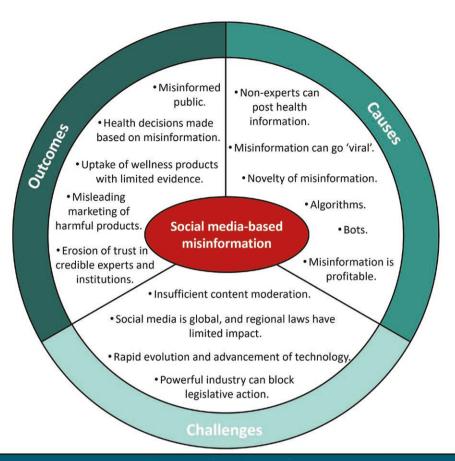
\*Corresponding author. Institute for Physical Acivity and Nutrition, School of Health and Social Development, Deakin University, 75 Pigdons Road, Geelong, Victoria 3216, Australia. E-mail: e.denniss@deakin.edu.au

### **Abstract**

Misinformation has been identified as a major threat to society and public health. Social media significantly contributes to the spread of misinformation and has a global reach. Health misinformation has a range of adverse outcomes, including influencing individuals' decisions (e.g. choosing not to vaccinate), and the erosion of trust in authoritative institutions. There are many interrelated causes of the misinformation problem, including the ability of non-experts to rapidly post information, the influence of bots and social media algorithms. Equally, the global nature of social media, limited commitment for action from social media giants, and rapid technological advancements hamper progress for improving information quality and accuracy in this setting. In short, it is a problem that requires a constellation of synergistic actions aimed at social media users, content creators, companies, and governments. A public health approach to social media-based misinformation that includes tertiary, secondary, and primary prevention may help address immediate impacts, long-term consequences, and root causes of misinformation. Tertiary prevention to 'treat' this problem involves increased monitoring, misinformation debunking, and warning labels on social media posts that are at a high risk of containing misinformation. Secondary prevention strategies include nudging interventions (e.g. prompts about preventing misinformation that appear when sharing content) and education to build media and information literacy. Finally, there is an urgent need for primary prevention, including systems-level changes to address key mechanisms of misinformation and international law to regulate the social media industry. Anything less means misinformation—and its societal consequences—will continue to spread.

Keywords: misinformation; social media; disinformation; health information; digital policy

Institute for Physical Activity and Nutrition, Deakin University, 75 Pigdons Road, Geelong, Victoria 3216, Australia



### **Prevention Strategies**

### **Tertiary**

- Monitor presence and impact of misinformation.
- Warning and fact-checking label interventions.
  - Refutation interventions involving debunking of misinformation.

### Secondary

- Strengthen resilience of individuals and communities.
- Nudging interventions to minimise spread.
- Educate to improve information and media literacy to enhance accuracy and sharing discernment.

### Primary

- Systems-level changes to address social media systems that drive misinformation.
- •International laws to regulate social media.

### Prioritise research into prevention strategies

### Contribution to Health Promotion

- Social media-based misinformation threatens public health through the provision of misleading information and undermines trust in credible experts and organizations.
- There are many interrelated causes of the misinformation problem and complex challenges for addressing it.
- A combination of complementary approaches to address misinformation and prevent its harms is required, informed by the World Health Organization competency framework for managing infodemics.
- We advocate for the uptake of widely recommended strategies, such as increased monitoring, misinformation debunking, and education interventions.
- We suggest international law as a potential strategy for bringing about systems-level changes to prevent misinformation.

### INTRODUCTION

We are living in the 'post-truth era'. The World Economic Forum's Global Risks Report identified misinformation and the technology that spreads it as a major global threat (World Economic Forum 2024). Furthermore, the World Health Organization (WHO) has made misinformation a priority area (WHO Regional Office for Europe 2022). Misinformation is defined by Treen et al. (2020) as 'misleading information that is created and spread, regardless of whether there is an intent to deceive', they also define disinformation as '[mis] information that is created and spread with intent to deceive'. Social media is now a useful tool for information dissemination in public health (De Vere Hunt and Linos 2022); however, it also contributes significantly to the spread of misinformation globally (Council of Canadian Academies [CCA] 2023). Social media-based misinformation is a complex problem with many consequences for public health (The Lancet 2025). This article will discuss the public health outcomes, causes, and challenges of misinformation. The evidence about intervention strategies is considered and using a public health prevention approach we suggest a combination of synergistic actions to address the immediate impacts, long-term consequences, and root causes of misinformation.

# PUBLIC HEALTH OUTCOMES OF MISINFORMATION VIA SOCIAL MEDIA

Health-related misinformation can mislead the public and thwart public health programmes. A large body of research has shown health misinformation, spanning a range of topics, including vaccines, infectious disease, nutrition, climate change, cancer, and smoking, is widely prevalent on major social media platforms (Wang et al. 2019, Treen et al. 2020, Suarez-Lledo and Alvarez-Galvez 2021, Denniss et al. 2023a). Even before the COVID-19 pandemic, there was an amplification of the 'anti-vaxx' movement, largely driven by the spread of misinformation, chiefly on social media (Hussain et al. 2018, Wang et al. 2019). Vaccine misinformation has led to reduced vaccination rates and outbreaks of disease, including measles in areas where elimination had been previously achieved (Hussain et al. 2018). The pandemic then brought vaccination to the forefront of public discourse. Global health misinformation networks published COVID-19 and vaccine misinformation, generating 3.8 billion views on Facebook in a 12-month period (Avaaz 2020). Misinformation and conspiracy theories contributed to COVID-19 vaccine hesitancy and caused increased fear and anxiety throughout the pandemic (Wilson and Wiysonge 2020, Rocha et al. 2023).

In addition, many so-called health and wellness influencers and brands use social media to spread misinformation for economic benefit (Lofft 2020, Baker 2022, CCA 2023). Recently, public health researchers have called for the social media industry to be recognized as a commercial determinant of health (CDoH), partly due to the financial incentives media giants receive to host misinformation on their platforms (Zenone et al. 2022). Marketing of products through sponsored advertising via payments to platforms and employment of influencers and brand presence is prolific across all major platforms, and marketing of health products, including supplements, weight loss, and fitness services, is particularly widespread (Lofft 2020, Denniss et al. 2023b). Research has shown wellness marketing content on social media often con-

tains misinformation (Baker 2022, CCA 2023, Denniss et al. 2024). The largely unregulated social media environment means that influencers and brands can continue to monetize their content while breaching the platform's community guidelines (Moran et al. 2024). Some health and wellness influencers have been known to cultivate alternative communities on social media and weaponize conspiracies, such as anti-vaxx misinformation, to sell wellness products (Baker 2022, Moran et al. 2024). These factors suggest that many social media users are exposed to misleading marketing content, which puts them at risk of purchasing health products and services with little to no evidence of efficacy, potentially leading to financial losses, minimal or no health benefits, and, in some instances, worse health outcomes.

Key industries whose systems, practices, and products are recognized as being CDoH also contribute to social mediabased misinformation. The tobacco industry, which has historically misled the public about the harms of tobacco products, has marketed new nicotine products (e.g. e-cigarettes or 'vapes' and smokeless tobacco) through social media influencers, some of whom claim the products are harmless (Tan and Bigman 2020). Tobacco and other harmful industries, such as the suntanning industry, are increasingly presenting themselves as scientific authorities by employing disinformation tactics, including emphasizing areas of scientific uncertainty and conducting or commissioning research that serves their interests (Reed et al. 2021). Furthermore, health professionals with financial conflicts of interest due to direct and/or indirect funding from the industry have published misinformation on social media. For example, in the USA, dietitians and doctors with large followings have partnered with big food companies to promote consumption of discretionary foods as part of the 'anti-diet' movement and assert that consumption of artificial sweeteners is not linked to any health risks, contradicting WHO guidance (Chavkin et al. 2023, 2024). Experimental research has found that young adults who were exposed to misleading social media content about vapes exhibited more favourable attitudes towards vapes compared with a control, suggesting that misleading marketing may be influential and a tactic that can be exploited by harmful industries (Albarracin et al. 2018).

Finally, the circumstances described above contribute to the erosion of trust in health and science-related fields. Health misinformation is often divisive, and wellness influencers are known to undermine public health and science authorities (Baker 2022, CCA 2023, Moran et al., 2024). Such healthrelated misinformation contributes to polarization and the public's mistrust in credible health experts and organizations (Penders 2018, CCA 2023, The Lancet 2025). Health professionals and experts with conflicts of interest, whether real or perceived, also contribute to the erosion of trust in public health (Penders et al. 2017). Many people instead trust alternative sources that spread misinformation and discredit legitimate health experts, creating a paradox: the mistrust fuelled by misinformation enables its continued spread (CCA 2023). This is problematic because mistrust has implications for the public's adherence to evidence-based health advice, which can impact the health of individuals and populations (CCA 2023). Once trust has been eroded, it is difficult and time-consuming to restore (Penders 2018, Caulfield 2020, CCA 2023). Furthermore, the prevalence of misinformation and erosion of trust in public health may also serve the interests of harmful

industries by undermining and distracting from the discussion of regulation and other measures.

### MECHANISMS THAT CAUSE MISINFORMATION

Social media environments are particularly conducive to the spread of misinformation, and there are several mechanisms that cause this. First, social platforms allow users to instantaneously publish on virtually any topic, regardless of their qualifications or information accuracy (Gajović and Svalastog 2016, Merchant and Asch 2018, World Economic Forum 2013). Platforms enable rapid sharing of content, and misinformation can 'go viral', cause harm, and change beliefs before it can be effectively corrected (World Economic Forum 2013, Merchant and Asch 2018, Vosoughi et al. 2018). Furthermore, research has shown that posts on X that contained misinformation received higher levels of engagement, spread more rapidly, and reached more users compared with posts that were truthful (Vosoughi et al. 2018). The same study also found that misinformation tended to be more novel, which is believed to be one reason it is spread more quickly than the truth (Vosoughi et al. 2018).

Second, social media algorithms contribute to the misinformation problem. Algorithms dictate content that appears in social media feeds based on users' previous online behaviour (Pariser 2011, Zimmer et al. 2019). They are a tool to individualize feeds to deliver users content they are likely to find interesting and engage with so they spend more time using the platform (Zenone et al. 2022). These sophisticated algorithms use individuals' social media data, including their social network, location, content they have viewed and engaged with, and content they have ignored, to predict what the user may prefer (Pariser 2011, Cohen 2018). Thus, social media users are shown an incomplete view of the scope of content published on a platform, whereby content that may challenge or oppose their beliefs is omitted (Pariser 2011, Cohen 2018). This can consolidate people's ideologies or beliefs because they may not be aware that opposing information is hidden. In some instances, algorithms have delivered users content that is incrementally more subversive and divisive, leading to radicalization and extreme views (Ribeiro et al. 2020, Matthews 2022). In the context of misinformation, the algorithm continues to feed misinformation to users who engage with it, and these users are unlikely to be presented with information that discredits or corrects it. Social media algorithms give rise to echo chambers (i.e. communities of people with similar views) (Bessi et al. 2015, 2016, Zimmer et al. 2019, Cinelli et al. 2021), and social network modelling has shown that misinformation spreads more rapidly within these conditions (Törnberg 2018).

Third, internet robots or 'bots' are known to automate the publication of misinformation and disinformation on social media. Social bots imitate human users on social media and automatically post and reshare content, often relating to controversial topics within politics and health (Yuan et al. 2019, Himelein-Wachowiak et al. 2021). Bots can operate together in a coordinated manner in a 'botnet' (Abokhodair et al. 2015) and use tactics, including tagging and replying to influential accounts to increase their exposure and manipulate users to reshare misinformation (Shao et al. 2018). Investigations by journalists have revealed the existence of groups that coordinate disinfor-

mation campaigns using social bots, some of whom have reportedly interfered with elections in multiple regions (Kirchgaessner et al. 2023, Hendy 2024). Bots can be difficult for social media users, researchers, and even social media platforms to identify, making them hard to counteract (Hendy 2024). However, tools for bot detection, e.g. Botometer, have been able to detect bot activity, which suggests social media giants could be doing more to identify and remove bots (Indiana University Observatory on Social Media 2024).

Fourth, content creators and social media platforms are financially incentivized for publishing misinformation. Content creators earn income from social media through user views, promoting their own products, and by partnering with brands who pay them to market products, with brands paying more to creators with greater followings and engagement (Hua et al. 2022, Simonson and Williams 2024). Influencers can make more money by creating content that generates higher views and engagement, and misinformation is linked to greater views and engagement, particularly when it is novel (Vosoughi et al. 2018). There is a risk, therefore, that creators may be financially rewarded for posting health misinformation (Moran et al. 2024). Simultaneously, platforms make greater profits from users spending more time on social media through advertising revenue and generating more user data, which is sold (Zenone et al. 2022). Platforms leverage algorithms to keep people online for longer and continue to deliver misinformation to users who are likely to engage with it. Furthermore, misinformation in and of itself can generate revenue because people spend time viewing it (Vosoughi et al. 2018). This phenomenon has been evident in relation to COVID-19 and vaccination, whereby influencers leveraged vaccine and other health misinformation to profit from engagement and product sales (Moran et al. 2024); meanwhile, social media giants profited from people spending time viewing COVID-19 misinformation (Avaaz 2020). The situation in which both the creator and platform are financially incentivized to publish misinformation is a vicious cycle that allows misinformation to flourish.

Lastly, an interplay between human psychology and algorithms contributes to misinformation's perpetuation (Rubin 2019). Most people are susceptible to confirmation bias, the inclination to believe information consistent with one's beliefs (Nickerson 1998), and repetition bias, the tendency to believe information one has heard multiple times (Hassan and Barber 2021). Algorithms repetitively present users with information that aligns with their beliefs, meaning it is likely to be believable. Furthermore, people are generally information overloaded and time poor and are therefore unable to comprehensively fact-check or critically evaluate the plethora of information they are presented with online (Rubin 2019). Health misinformation can be difficult to detect, and many individuals have limited information, media, and health literacy, which makes it more difficult to discern accuracy (Rubin 2019, Australian Institute of Health and Welfare 2023). Humans also tend to exhibit truth bias, which makes them inclined to believe that the information they are presented with is true unless given significant reason to believe otherwise (Levine 2014, Rubin 2019). In combination, these factors contribute to humans' susceptibility to misinformation, and research has shown that humans are more likely than bots to repost misinformation (Vosoughi et al. 2018).

## CHALLENGES FOR CURBING MISINFORMATION VIA SOCIAL MEDIA

Social media-based misinformation is a complex problem. There are numerous complicated challenges for addressing the mechanisms described above and the broader commercial and societal context within which this public health problem exists. One challenge is that social media operates globally, which creates complexity for legislation. Typically, legislative action is taken by a country or region (Centre for Digital Wellbeing 2021) and is therefore likely to have a limited impact. For example, an Australian regulatory body banned influencers from posting testimonials about therapeutic goods, including sunscreen and supplements (Therapeutic Goods Administration [TGA] 2022a), and fined individuals and businesses for posting inaccurate information about therapeutic goods (TGA 2021, 2022b). However, Australian social media users are likely to be exposed to material from international influencers who are not obligated to abide by Australian laws. Comprehensive monitoring of social media content for compliance with regulations is also difficult to achieve due to the volume of content posted on social media daily. Furthermore, this type of regulation only penalizes the entities responsible for publishing misleading information, and the platforms that facilitate misinformation face no recourse.

An additional challenge is that social media giants have shown insufficient commitment to content moderation and the removal of misinformation. Thus far, actions taken by platforms have been shown to be limited in their effectiveness. For example, the analysis of misinformation about the COVID-19 pandemic showed that only 16% of misinformation on Facebook carried a warning label (Avaaz 2020). A former Meta employee and whistleblower revealed research documents from Meta that indicate the company is aware that misinformation, divisive content, and hate speech on their platforms impact societies globally, and only a small proportion of this content is removed (Pelley 2021). Elon Musk, who acquired X in 2022, reduced the size of his content moderation team and has publicly voiced his aversion to content moderation (CBS News 2022). In February 2023, Musk posted on X, 'All things in moderation, especially content moderation' (Musk 2023). More recently, in January 2025, Mark Zuckerberg announced that Meta would be dropping their independent third-party fact-checkers, stating that they led to 'too much censorship' without effectively addressing misinformation (Watt et al. 2025), despite evidence that fact-checking can reduce belief in misinformation (Porter and Wood 2021). These examples demonstrate the lack of commitment to meaningful content moderation by major platforms, which may be due to the large investment that would be required and the loss of revenue from people viewing divisive content.

A further barrier is that technology and social media are continuously and rapidly evolving. Platforms can make seemingly instantaneous changes to their algorithm or interface that have significant implications for the functionality of the platform for millions of users. There has been a surge in artificial intelligence (AI) in recent years, much of which has been funded by tech giants, including Meta and Google (Buttazzo 2023). AI is predicted to reach human intelligence by 2030 and far exceed it in the years that follow (Buttazzo 2023), and social media platforms, including Instagram, LinkedIn, and Snapchat, have embedded AI tools into their platforms. Many

implications of technological innovations are not known until after they have been rolled out, and it can then be difficult to put the genie back into the bottle. These rapid and unprecedented technological advancements make it difficult to develop and implement timely and effective policies. Meta CEO Mark Zuckerberg's motto, 'move fast and break things' (Hamilton 2022), suggests big players in the industry not only understand but also take advantage of this.

Finally, social media companies are powerful corporate entities with extensive resources (Rabb 2018). Historically, these companies have circumvented consequences for unethical practice with relative ease (Newcomb 2018, McCallum 2022). In 2023, Meta, YouTube, and TikTok's revenue totalled \$134 billion, \$32 billion, and \$120 billion USD, respectively (Statista 2024a, 2024b, WARC 2024). Meta's responses to privacy breaches on Facebook have historically involved apologies to users but limited changes to policy or operations (Newcomb 2018). After the 2018 Cambridge Analytica scandal, whereby 87 million Facebook users' private data was harvested from Facebook and used for political advertising, Meta agreed to pay a \$725 million USD settlement, a small percentage of their annual revenue (McCallum 2022). Since the scandal, Meta has continued to generate massive revenue and maintain a large user base. Changes to legislation that impact social media have resulted in backlash. In 2021, the Australian government passed a law that mandated social platforms pay news media outlets for publishing their articles (Meade 2021). Facebook reacted by blocking Australian users from accessing news content, which also blocked the pages of Australian health and emergency services, resulting in a key communication channel being disrupted. Meta leveraged this to negotiate changes to the legislation before unblocking pages and news content (Meade 2021). This example demonstrates the power that social media companies wield to block government actions and the impact their backlash can have on the people and organizations relying on their services.

# A WAY FORWARD: STRATEGIES FOR MISINFORMATION PREVENTION

Thus far, regulations and voluntary actions from social media companies have had a limited impact in addressing the misinformation problem, and urgent action is required. As has been seen with the regulation of other harmful industries, such as the tobacco industry, curbing the activities of social media companies is likely to take time, be fiercely opposed, and be difficult to implement (Daube and Maddox 2021). The literature suggests that complex problems are not 'solved' per se but can be managed and addressed through a range of incremental and complementary approaches (Head 2019). Therefore, a combination of priority strategies is suggested below. In public health, there is an emphasis on holistic approaches that involve tertiary, secondary, and primary prevention to address the immediate acute impacts, long-term consequences, and root causes of health issues (Goldsteen et al. 2010). Social media-based misinformation is a public health issue with immediate impacts (e.g. a misinformed public), long-term consequences (e.g. eroded trust in science and public health), and complex root causes (e.g. algorithms, bots, and financial interests). As such, a public health prevention approach to managing and addressing misinformation may be beneficial, and a combination of tertiary, secondary, and primary prevention strategies is presented.

### Tertiary prevention

Misinformation is widespread on social media, and tertiary prevention strategies to respond to and 'treat' the problem are required. The WHO competency framework for managing infodemics provides guidance for public health organizations in responding to misinformation. Whilst it was designed in the context of public health emergencies, many of the strategies remain relevant. The WHO framework has four domains, three of which relate to tertiary prevention. First, measure and monitor the impact of misinformation; second, determine how it is spread; and third, intervene to minimize consequences (Rubinelli et al. 2022). Interventions involving the inclusion of warning, fact-checking, and source-credibility labels on high-risk information on social media are effective strategies for minimizing harm (Kozyreva et al. 2024). Specifically, warning labels are used to caution users that a post is likely to contain misinformation, and fact-checking labels specify that a post contains false or partially false information. There is evidence that warning and fact-checking labels can improve accuracy, discernment, and sharing intentions (Kozyreva et al. 2024). During the COVID-19 pandemic, there was pressure on social media giants to prevent misinformation, and major platforms (e.g. Facebook and YouTube) placed warning labels that provided links to credible information (e.g. WHO website) on posts that discussed COVID-19 (Warnke et al. 2024). Advocates should put pressure on social media giants to embed warning and fact-checking labels in their platforms and provide access to their application programming interfaces to allow governments and public health organizations to more easily and comprehensively monitor misinformation trends.

Refutation interventions involving debunking of misinformation, whereby corrective information is provided, are an evidence-based strategy that has been shown to reduce belief in rumours and misinformation (Danielson et al. 2024, Kozyreva et al. 2024). While there has been concern about the possibility of debunking to backfire and reinforce misconceptions, a recent meta-analysis found no evidence of a backfiring effect and recommended the refutation of misinformation is published online (Danielson et al. 2024). Public health advocates and the broader health community, including medical professionals, academics, scientists, and public health organizations, have an important role to play in refuting misinformation and promoting evidence-based information in both online and offline settings, and this has been encouraged by others in the field (Ecker et al. 2024). While doctors and other health professionals have been seen to contribute to misinformation in some instances, they are generally well trusted by the community and can contribute to misinformation refutation in clinical and community settings, as was seen during the COVID-19 pandemic in rural areas of the USA (Howard 2021, Mainous et al. 2024).

#### Secondary prevention

The final domain of the WHO framework relates to secondary prevention and involves strengthening the resilience of individuals and communities to misinformation (Rubinelli et al. 2022). There is a growing body of evidence investigating the impact of interventions on misinformation detection and reducing the sharing of misinformation (Kozyreva et al. 2024). 'Nudging' or prompting strategies have been used as components of comprehensive approaches to health promo-

tion for many years, and research has shown that nudging can improve sharing intentions and discernment (Kozyreva et al. 2024). Examples of nudging include messages that remind people about the harms of misinformation or prompts to read an article in full and evaluate the accuracy of the headline before sharing (Kozyreva et al. 2024). Previously, X (then Twitter) embedded a 'read before you retweet' nudge in the platform that reminded users to read articles when reposting links that they had not clicked on (Verge 2020). Similar nudges could be embedded in social media platforms as a strategy to minimize the spread of misinformation.

Educational interventions to improve digital and media literacy have also shown promise in improving individuals' accuracy, credibility, and sharing discernment (Kozyreva et al. 2024). Digital and media literacy skills are now vital life skills and competencies that are required at every life stage and within almost every vocation. School curricula for children are an obvious setting, and evidence-based programs show effectiveness (Dumitru et al. 2022). The need for information and media literacy education has been echoed by the WHO and others in the field (Rubin 2019, WHO Regional Office for Europe 2022), and strengthening the information and media literacy of populations is a priority area for reducing the spread and impact of misinformation. There is evidence that when the companies or industries involved in a public health problem contribute to education programs, the educative content has been skewed, e.g. stealth marketing of food brands in children's food and nutrition education in schools (Wilkinson 2024) and industry-funded youth education programs about alcohol, tobacco and gambling that serve industry interests (Sebrié and Glantz 2007, van Schalkwyk et al. 2022a, Van Schalkwyk et al. 2022b). Therefore, education programs should be developed and run by governments, independent health agencies or other groups without conflicts of interest.

### Primary prevention

The adage 'a lie can travel halfway around the world before the truth can get its boots on' speaks to the need for primary prevention of misinformation. It is acknowledged in the literature that in addition to individual- and community-level interventions, systems-level changes are required to address the misinformation problem (Kozyreva et al. 2024). The current social media system not only facilitates the publication and spread of misinformation but also fuels and generates profits from it. Historically, voluntary corporate social responsibility actions by commercial actors to promote public health or prevent harm have been shown to hinder or delay meaningful and legitimate actions (Madureira Lima and Galea 2018, De Lacy-Vawdon and Livingstone 2020). Due to the highly profitable nature of social media, impactful voluntary action is unlikely; therefore, greater regulation is necessary. Many governments have policies that aim to regulate various facets of social media and minimize misinformation, including the European Union's Digital Services Act, which came into effect throughout 2023-2024 (Centre for Digital Wellbeing 2021, European Union 2022). The Australian government recently proposed new laws to penalize social media companies for mis- and disinformation on their platforms; however, this was met with a mix of support and criticism and did not pass in the Senate (Rownland 2024, Speers and Truu 2024). Despite the challenges that exist, it is clear there is potential and an appetite in some jurisdictions for greater regulation.

The global nature of social media is a barrier to policy and legislative action. In an increasingly globalized world, the need for global health law to address widespread public health problems has been recognized, and treaty law is considered a promising tool to facilitate cooperation between states to achieve public health objectives that cannot be adequately addressed through regional or domestic laws and policies (Taylor 2017). One example is the WHO Framework Convention on Tobacco Control (FCTC), which is a treaty subject to international law that came into force in 2005 as a means for countries to cooperate to prevent the harms of tobacco (McInerney 2019). States that signed the FCTC agreed to actions including increased taxes, banning advertising, and regulating the content of tobacco products. Analysis of the FCTC found evidence that it had positively impacted tobacco control, in part due to its status as an international legal obligation (McInerney 2019). Links have been drawn between social media and tobacco because both are purposefully addictive and sold by powerful profit-driven companies (MacBride 2018, Holly et al. 2024). Therefore, the success of the FCTC in addressing the impacts of tobacco suggests that international law to regulate social media may be a promising strategy to address misinformation. Regulatory action via treaty could include mandating greater content moderation, removal of bots, and changes to algorithms to present users with a balanced view of content and show results from credible authorities when platforms' search engines are used, similar to action taken by Pinterest for vaccine-related searches (Ozoma 2019).

It is important to note that international agreement on treaty law will be difficult to achieve. Regulation of harmful industries is challenging to implement at the national level, and these challenges are likely to be magnified at the international level due to diverging national interests (Taylor 2017). Negotiating international law is complex, which has historically made commitment and implementation by member states a time-consuming process (Taylor 2017). Furthermore, harmful industries have learned from the tobacco experience and use innovative tactics to avoid regulation and protect their profits (Thomas et al. 2024). It is therefore important that other avenues of legislation and systems-level change are also explored and targeted. To the knowledge of the authors, most misinformation research has focused on tertiary and secondary prevention and the impact of primary prevention; systems-level strategies for addressing misinformation are an important research gap that should be addressed to inform future policies and interventions.

### **CONCLUSION**

In conclusion, a combination of synergistic primary, secondary, and tertiary prevention strategies is required to address the public health impacts of social media-based misinformation. This perspective piece advocates for the use of the WHO framework for managing infodemics in conjunction with established and well-evidenced actions, such as monitoring, debunking, warning labels, and education. We contribute to the field by suggesting international treaty law as a primary prevention strategy to bring about the systems-level changes required to address the mechanisms and root causes of misinformation. International regulation and systems-level changes may also serve to address the broader health and societal impacts of social media, such as its purposefully addictive

features. Meaningful cooperation from social media giants in all levels of prevention of misinformation would lead to a greater impact. Wherever possible, public health practitioners, researchers, organizations, and the broader health community should advocate for greater action from the social media industry to address misinformation. Finally, primary prevention and systems-level strategies to address misinformation should be a research priority area.

### **ACKNOWLEDGEMENTS**

The authors acknowledge Prof. Sarah A. McNaughton for her contributions to the PhD studies that helped seed this article.

### CONFLICT OF INTEREST

None declared.

### **FUNDING**

None declared.

### **DATA AVAILABILITY**

No new data were generated or analysed in support of this research.

### **ETHICAL APPROVAL**

No new data were generated or analysed in support of this research, and approval from an ethics committee was therefore not required.

### **AUTHOR CONTRIBUTIONS**

E.D. generated the initial idea for this manuscript. E.D. and R.L. sourced relevant literature and together generated the ideas for the suggested strategies included in the manuscript. E.D. wrote the first draft, and R.L. critically revised the manuscript for important intellectual content.

### **REFERENCES**

Abokhodair N, Yoo D, McDonald DW. Dissecting a social botnet: growth, content and influence in Twitter. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. Vancouver, BC, Canada: Association for Computing Machinery, 2015.

Albarracin D, Romer D, Jones C *et al*. Misleading claims about tobacco products in YouTube videos: experimental effects of misinformation on unhealthy attitudes. *J Med Internet Res* 2018;**20**:e229. https://doi.org/10.2196/jmir.9959

Australian Institute of Health and Welfare. 2023. *Australia's Health* 2018, *Health Literacy*. https://www.aihw.gov.au/reports/australias-health/australias-health-2018/contents/indicators-of-australias-health/health-literacy (19 July, date last accessed).

Avaaz. Facebook's Algorithm: A Major Threat to Public Health. New York: Avaaz, 2020.

Baker SA. Alt. Health Influencers: how wellness culture and web culture have been weaponised to promote conspiracy theories and far-right extremism during the COVID-19 pandemic. *Eur J Cult Stud* 2022;25:3–24. https://doi.org/10.1177/13675494211062623

Bessi A, Coletto M, Davidescu GA et al. Science vs conspiracy: collective narratives in the age of misinformation. PLoS One

- 2015;10:e0118093–17. https://doi.org/10.1371/journal.pone.0118093
- Bessi A, Petroni F, Vicario MD *et al*. Homophily and polarization in the age of misinformation. *Eur Phys J Spec Top* 2016;225:2047–59. https://doi.org/10.1140/epist/e2015-50319-0
- Buttazzo G. Rise of artificial general intelligence: risks and opportunities. Front Artif Intell 2023;6:1226990. https://doi.org/10.3389/frai.2023.1226990
- Caulfield T. *The COVID-19 Pandemic Will Cause Trust in Science to Be Irreparably Harmed*. 2020. https://www.theglobeandmail.com/opinion/article-the-covid-19-pandemic-will-cause-trust-in-science-to-be-irreparably/ (3 July 2024, date last accessed).
- CBS News. Musk Fires Outsourced Content Moderators Who Track Abuse on Twitter. 2022. https://www.cbsnews.com/news/elon-musk-twitter-layoffs-outsourced-content-moderators/ (9 July, date last accessed).
- Centre for Digital Wellbeing. *International Regulation of Social Media*. Kingston: Centre for Digital Wellbeing, 2021.
- Chavkin S, Gilbert C, O'Connor A. US Regulator Cracks Down on Food Industry for Paid Dietitian 'Influencer' Posts. 2023. https://www.theexamination.org/articles/ftc-crackdown (9 December 2024, date last accessed).
- Chavkin S, Tsui A, Gilbert C et al. As Obesity Rises, Big Food and Dietitians Push 'Anti-Diet' Advice. 2024. https://www.theexamination.org/articles/as-obesity-rises-big-food-and-dietitians-push-anti-diet-advice (9 December 2024, date last accessed).
- Cinelli M, De Francisci Morales G, Galeazzi A *et al.* The echo chamber effect on social media. *Proc Natl Acad Sci USA* 2021;118:e2023301118. https://doi.org/10.1073/pnas.2023301118
- Cohen JN. Exploring echo-systems: how algorithms shape immersive media environments. *J Media Lit Educ* 2018;10:139–51. https://doi.org/10.23860/jmle-2018-10-2-8
- Council of Canadian Academies. Fault Lines. Expert Panel on the Socioeconomic Impacts of Science and Health Misinformation. Ottawa: Council of Canadian Academies, 2023.
- Danielson RW, Jacobson NG, Patall EA *et al.* The effectiveness of refutation text in confronting scientific misconceptions: a meta-analysis. *Educ Psychol* 2024;60:1–25. https://doi.org/10.1080/00461520.2024.2365628
- Daube M, Maddox R. Impossible until implemented: New Zealand shows the way. *Tob Control* 2021;30:361–2. https://doi.org/10.1136/tobaccocontrol-2021-056776
- De Lacy-Vawdon C, Livingstone C. Defining the commercial determinants of health: a systematic review. *BMC Public Health* 2020;20:1022. https://doi.org/10.1186/s12889-020-09126-1
- De Vere Hunt I, Linos E. Social media for public health: framework for social media–based public health campaigns. *J Med Internet Res* 2022;24:e42179. https://doi.org/10.2196/42179
- Denniss E, Lindberg R, Marchese LE et al. #Fail: the quality and accuracy of nutrition-related information by influential Australian Instagram accounts. Int J Behav Nutr Phys Act 2024;21:16. https://doi.org/10.1186/s12966-024-01565-y
- Denniss E, Lindberg R, McNaughton SA. Nutrition-related information on Instagram: a content analysis of posts by popular Australian accounts. *Nutrients* 2023b;15:2332. https://doi.org/10.3390/nu15102332
- Denniss E, Lindberg R, McNaughton SA. Quality and accuracy of online nutrition-related information: a systematic review of content analysis studies. *Public Health Nutr* 2023a;26:1345–57. https://doi.org/10.1017/s1368980023000873
- Dumitru E-A, Ivan L, Loos E. A Generational Approach to Fight Fake News: In Search of Effective Media Literacy Training and Interventions. Switzerland: Springer International Publishing, 2022, 291–310.
- Ecker U, Roozenbeek J, Van Der Linden S et al. Misinformation poses a bigger threat to democracy than you might think. Nature 2024;630:29–32. https://doi.org/10.1038/d41586-024-01587-3
- European Union. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market

- For Digital Services and amending Directive 2000/31/EC (Digital Services Act). 2022.
- Gajović S, Svalastog AL. When communicating health-related knowledge, beware of the black holes of the knowledge landscapes geography. *Croat Med J* 2016;57:504–9. https://doi.org/10.3325/cmj.2016.57.504
- Goldsteen RL, Goldsteen K, Graham DC. Introduction to Public Health. New York: Springer Publishing Company, Incorporated, 2010
- Hamilton IA. Mark Zuckerberg Hasn't Let Go of 'Move Gast and Break Things'. 2022. https://www.businessinsider.com/meta-mark-zuckerberg-new-values-move-fast-and-break-things-2022-2 (9 December 2024, date last accessed).
- Hassan A, Barber SJ. The effects of repetition frequency on the illusory truth effect. Cogn Res 2021;6. https://doi.org/10.1186/s41235-021-00301-5
- Head BW. Forty years of wicked problems literature: forging closer links to policy studies. *Policy Soc* 2019;38:180–97. https://doi.org/ 10.1080/14494035.2018.1488797
- Hendy E. Who Trolled Amber? Inside the Podcast Exposing the Horrors of the Depp/Heard Trial. 2024. https://www.independent.co.uk/life-style/amber-heard-trolling-johnny-depp-trial-b2509469. html (22 July, date last accessed).
- Himelein-Wachowiak M, Giorgi S, Devoto A *et al.* Bots and misinformation spread on social media: implications for COVID-19. *J Med Internet Res* 2021;23:e26933. https://doi.org/10.2196/26933
- Holly L, Demaio S, Kickbusch I. Public health interventions to address digital determinants of children's health and wellbeing. *Lancet Public Health* 2024;9:e700–4. https://doi.org/10.1016/s2468-2667(24)00180-4
- Howard B. *Talking to Vaccine Skeptics in Rural, Conservative America*. Washington: Association of American Medical Colleges, 2021.
- Hua Y, Horta Ribeiro M, Ristenpart T et al. Characterizing alternative monetization strategies on YouTube. Proc ACM Hum-Comput Interact 2022;6:1–30. https://doi.org/10.1145/3555174
- Hussain A, Ali S, Ahmed M *et al.* The anti-vaccination movement: a regression in modern medicine. *Cureus* 2018;10:e2919. https://doi.org/10.7759/cureus.2919
- Indiana University Observatory on Social Media. *Introducing Botometer X*. 2024. https://osome.iu.edu/research/blog/introducing-botometer-x (22 July, date last accessed).
- Kirchgaessner S, Ganguly M, Pegg D et al. Revealed: The Hacking and Disinformation Team Meddling in Elections. 2023. https://www.theguardian.com/world/2023/feb/15/revealed-disinformation-team-jorge-claim-meddling-elections-tal-hanan (22 July, date last accessed).
- Kozyreva A, Lorenz-Spreen P, Herzog SM et al. Toolbox of individual-level interventions against online misinformation. Nat Hum Behav 2024;8:1044–52. https://doi.org/10.1038/s41562-024-01881-0
- The Lancet. Health in the age of disinformation. *Lancet* 2025;405:173. https://doi.org/10.1016/s0140-6736(25)00094-7
- Levine TR. Truth-default theory (TDT): a theory of human deception and deception detection. *J Lang Soc Psychol* 2014;33:378–92. https://doi.org/10.1177/0261927x14535916
- Lofft Z. When social media met nutrition. Health Sci Inquiry 2020;11:56–61. https://doi.org/10.29173/hsi319
- MacBride E. Is Social Media the Tobacco Industry of the 21st Century? 2018. https://www.forbes.com/sites/elizabethmacbride/2017/12/31/is-social-media-the-tobacco-industry-of-the-21st-century/ (26 August, date last accessed).
- Madureira Lima J, Galea S. Corporate practices and health: a framework and mechanisms. *Glob Health* 2018;14:21. https://doi.org/10.1186/s12992-018-0336-y
- Mainous AG, Sharma P, Yin L et al. Conflict among experts in health recommendations and corresponding public trust in health experts. Front Med 2024;11:1430263. https://doi.org/10.3389/fmed.2024.1430263
- Matthews J. Radicalization Pipelines: How Targeted Advertising on Social Media Drives People to Extremes. 2022. https://theconversation.

- com/radicalization-pipelines-how-targeted-advertising-on-so-cial-media-drives-people-to-extremes-173568 (4 September 2023, date last accessed).
- McCallum S. Meta Settles Cambridge Analytica Scandal Case for \$725m. 2022. https://www.bbc.com/news/technology-64075067 (16 Iuly, date last accessed).
- McInerney TF. The WHO FCTC and global governance: effects and implications for future global public health instruments. *Tob Control* 2019;28:s89–93. https://doi.org/10.1136/tobaccocontrol-2018-054358
- Meade A. Facebook Reverses Australia News Ban After Government Makes Media Code Amendments. 2021. https://www.theguardian.com/media/2021/feb/23/facebook-reverses-australia-news-ban-after-government-makes-media-code-amendments (3 October 2023, date last accessed).
- Merchant RM, Asch DA. Protecting the value of medical science in the age of social media and 'Fake News'. *J Am Med Assoc* 2018;320:2415–6. https://doi.org/10.1001/jama.2018.
- Moran RE, Swan AL, Agajanian T. Vaccine misinformation for profit: conspiratorial wellness influencers and the monetization of alternative health. *Int J Commun* 2024;18:1202–24.
- Musk E. All Things in Moderation, Especially Content Moderation. In @elonmusk. 2023. https://x.com/elonmusk/status/1627377252777 226242?lang=en.X (9 July 2024, date last accessed).
- Newcomb A. A Timeline of Facebook's Privacy Issues—and Its Responses. 2018. https://www.nbcnews.com/tech/social-media/ timeline-facebook-s-privacy-issues-its-responses-n859651 (15 July, date last accessed).
- Nickerson RS. Confirmation bias: a ubiquitous phenomenon in many guises. Rev Gen Psychol 1998;2:175–220. https://doi.org/10.1037//1089-2680.2.2.175
- Ozoma I. Bringing Authoritative Vaccine Results to Pinterest Search. 2019. https://newsroom-archive.pinterest.com/en-gb/bringing-authoritative-vaccine-results-to-pinterest-search (2 October 2023, date last accessed).
- Pariser E. The Filter Bubble: What the Internet Is Hiding from You. New York: Penguin, 2011.
- Pelley S. Whistleblower: Facebook Is Misleading the Public on Progress against Hate Speech, Violence, Misinformation. 2021. https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021-10-03/ (28 November 2023, date last accessed).
- Penders B. Why public dismissal of nutrition science makes sense: post-truth, public accountability and dietary credibility. *Br Food J Croydon, England*) 2018;**120**:1953–64. https://doi.org/10.1108/BFJ-10-2017-0558
- Penders B, Wolters A, Feskens E *et al.* Capable and credible? Challenging nutrition science. *Eur J Nutr* 2017;56:2009–12. https://doi.org/10.1007/s00394-017-1507-y
- Porter E, Wood TJ. The global effectiveness of fact-checking: evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. *Proc Natl Acad Sci USA* 2021;118:e2104235118. https://doi.org/10.1073/pnas.2104235118
- Rabb N. How Powerful Are Social Media Companies? 2018. https:// ricknabb.medium.com/how-powerful-are-social-media-companies-4d1a3f6c6b3c (9 December 2024, date last accessed).
- Reed G, Hendlin Y, Desikan A *et al.* The disinformation playbook: how industry manipulates the science-policy process—and how to restore scientific integrity. *J Public Health Policy* 2021;42:622–34. https://doi.org/10.1057/s41271-021-00318-6
- Ribeiro MH, Ottoni R, West R *et al.* Auditing radicalization pathways on YouTube. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, 2020. https://doi.org/10.1145/3351095.3372879
- Rocha YM, De Moura GA, Desidério GA *et al*. The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review. *J Public Health* 2023;31:1007–16. https://doi.org/10.1007/s10389-021-01658-z

- Rownland M. Communications Legislation Amendment (Combatting Misinformation and Disinformation) Bill 2024. Canberra: Australian Government, 2024.
- Rubin VL. Disinformation and misinformation triangle: a conceptual model for 'fake news' epidemic, causal factors and interventions. *I Doc* 2019;75:1013–34. https://doi.org/10.1108/id-12-2018-0209
- Rubinelli S, Purnat TD, Wilhelm E *et al.* WHO competency framework for health authorities and institutions to manage infodemics: its development and features. *Hum Res Health* 2022;20:49. https://doi.org/10.1186/s12960-022-00750-z
- Sebrié EM, Glantz SA. Attempts to undermine tobacco control. *Am J Public Health* 2007;97:1357–67. https://doi.org/10.2105/AJPH.2006.094128
- Shao C, Ciampaglia GL, Varol O *et al*. The spread of low-credibility content by social bots. *Nat Commun* 2018;9:4787. https://doi.org/10.1038/s41467-018-06930-7
- Simonson J, Williams R. How To Make Money On Social Media in 2024. 2024. https://www.forbes.com/advisor/business/how-make-money-social-media/#:~:text=Social%20media%20monetization%20refers%20to,collaborations%2C%20sponsored%20content%20and%20advertising (22 July, date last accessed).
- Speers D, Truu M. Media watchdog will have more power to force tech companies to crack down on disinformation under new bill. 2024. https://www.abc.net.au/news/2024-09-11/acmacrackdown-social-media-disinformation-bill/104339144? utm\_medium=social&utm\_content=sf274604406&utm\_campaign=tw\_abc\_news&utm\_source=t.co&sf274604406=1 (12 December 2024, date last accessed).
- Statista. Annual Revenue and Net Income Generated by Meta Platforms from 2007 to 2023. 2024a. https://www.statista.com/statistics/277229/facebooks-annual-revenue-and-net-income/#:~:text=In%202023%2C%20Meta%20Platforms%20generated,119%20billion%20USD%20in%202022. (11 July, date last accessed).
- Statista. Worldwide Advertising Revenues of YouTube as of 1st Quarter 2024. 2024b. https://www.statista.com/statistics/289657/youtube-global-quarterly-advertising-revenues/ (11 July, date last accessed).
- Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res* 2021;23:e17187. https://doi.org/10.2196/17187
- Tan ASL, Bigman CA. Misinformation about commercial tobacco products on social media—implications and research opportunities for reducing tobacco-related health disparities. *Am J Public Health* 2020;110:S281–3. https://doi.org/10.2105/AJPH.2020.305910
- Taylor AL. Global health law: international law and public health policy. In: Quah SR (ed.), *International Encyclopedia of Public Health* (2nd edn). Oxford: Academic Press, 2017, 268–81.
- Therapeutic Goods Administration. *Peter Evans Chef Pty Ltd Fined* \$79,920 for Alleged Unlawful Advertising. 2021. https://www.tga.gov.au/news/media-releases/peter-evans-chef-pty-ltd-fined-79920-alleged-unlawful-advertising (28 August, date last accessed).
- Therapeutic Goods Administration. JSHealth Vitamins Pty Ltd Fined \$26,640 for Alleged Unlawful Advertising. 2022b. https://www.tga.gov.au/news/media-releases/jshealth-vitamins-pty-ltd-fined-26640-alleged-unlawful-advertising (25 November 2022, date last accessed).
- Therapeutic Goods Administration. TGA Social Media Advertising Guide. 2022a. https://www.tga.gov.au/resources/resource/guidance/tga-social-media-advertising-guide#social (2 October 2023, date last accessed).
- Thomas S, Daube M, Van Schalkwyk M *et al.* Acting on the commercial determinants of health. *Health Promot Int* 2024;39. https://doi.org/10.1093/heapro/daae183
- Törnberg P. Echo chambers and viral misinformation: modeling fake news as complex contagion. *PLoS One* 2018;13:e0203958–21. https://doi.org/10.1371/journal.pone.0203958
- Treen KMDI, Williams HTP, O'Neill SJ. Online misinformation about climate change. WIREs Clim Change 2020;11. https://doi.org/10.1002/wcc.665

van Schalkwyk MCI, Hawkins B, Petticrew M. The politics and fantasy of the gambling education discourse: an analysis of gambling industry-funded youth education programmes in the United Kingdom. SSM - Popul Health 2022a;18:101122. https://doi.org/10.1016/j.ssmph.2022.101122

- Van Schalkwyk MCI, Petticrew M, Maani N *et al.* Distilling the curriculum: an analysis of alcohol industry-funded school-based youth education programmes. *PLoS One* 2022b;17:e0259560. https://doi.org/10.1371/journal.pone.0259560
- Verge V. Twitter Is Bringing Its 'Read Before You Retweet' Prompt to All Users. 2020. https://www.theverge.com/2020/9/25/21455635/twitter-read-before-you-tweet-article-prompt-rolling-out-globally-soon (11 December 2024, date last accessed).
- Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018;359:1146-51. https://doi.org/10.1126/science.aap9559
- Wang Y, McKee M, Torbica A et al. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019;240:112552. https://doi.org/10.1016/j. socscimed.2019.112552
- WARC. Reports Suggest TikTok Revenues up 40% to \$120bn in 2023. 2024. https://www.warc.com/content/feed/reports-suggest-tiktok-revenues-up-40-to-120bn-in-2023/en-GB/9285#:~:text=Short% 20video%20platform%20TikTok%20is,continue%20operating%20in%20the%20country (11 July, date last accessed).
- Warnke L, Maier A-L, Gilbert DU. Social media platforms' responses to COVID-19-related mis- and disinformation: the insufficiency of self-governance. J Manage Gov 2024;28:1079–115. https://doi. org/10.1007/s10997-023-09694-5

- Watt N, Riedlinger M, Montaña-Niño S. Meta Is Abandoning Fact Checking This Doesn't Bode Well for the Fight Against Misinformation. 2025. https://theconversation.com/meta-is-abandoning-fact-checking-this-doesnt-bode-well-for-the-fight-against-misinformation-246878 (20 January 2025, date last accessed).
- Wilkinson E. Food industry has infiltrated UK children's education: stealth marketing exposed. *BMJ* 2024;387:q2661. https://doi.org/10.1136/bmj.q2661
- Wilson SL, Wiysonge C. Social media and vaccine hesitancy. BMJ Glob Health 2020;5:e004206. https://doi.org/10.1136/bmjgh-2020-004206
- World Economic Forum. *Global Risks* 2013. Geneva, Switzerland: World Economic Forum, 2013.
- World Economic Forum. *Global Risks Report* 2024. Geneva, Switzerland: World Economic Forum, 2024.
- World Health Organization Regional Office for Europe. Toolkit for Tackling Misinformation on Noncommunicable Diseases. Copenhagen: World Health Organization Regional Office for Europe, 2022.
- Yuan X, Schuchard RJ, Crooks AT. Examining emergent communities and social bots within the polarized online vaccination debate in Twitter. Soc Media Soc 2019;5:205630511986546. https://doi. org/10.1177/2056305119865465
- Zenone M, Kenworthy N, Maani N. The social media industry as a commercial determinant of health. *Int J Health Policy Manag* 2022. https://doi.org/10.34172/ijhpm.2022.6840
- Zimmer F, Scheibe K, Stock M *et al.* Fake news in social media: bad algorithms or biased users? *J Inf Sci Theory Pract* 2019;7:40–53. https://doi.org/10.1633/JISTaP.2019.7.2.4