BMJ Open Impact of strategies to mitigate misinformation in diverse settings and populations: a protocol for a living evidence synthesis

Michael Wilson , , Marcela Vélez, John Lavis, , Misinformation Living Evidence Synthesis Team

To cite: Wilson M. Vélez M. Lavis J. et al. Impact of strategies to mitigate misinformation in diverse settings and populations: a protocol for a living evidence synthesis. BMJ Open 2023;13:e076672. doi:10.1136/ bmjopen-2023-076672

Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/ bmjopen-2023-076672).

Received 13 June 2023 Accepted 29 September 2023

Check for updates

@ Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by

¹Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Ontario, Canada

²McMaster Health Forum, McMaster University, Hamilton, Ontario, Canada

Correspondence to

Dr Michael Wilson: wilsom2@mcmaster.ca

ABSTRACT

Introduction Misinformation refers to inadvertent misleading information that the public may be exposed and share without intent to cause harm, and can delay or prevent effective care, affect mental health, lead to misallocation of health resources and/or create or exacerbate public-health crises. There are many strategies to address misinformation, but there is a need to evaluate their effects. Our objective is to synthesise and routinely update evidence to assess the impact of strategies to mitigate health-related misinformation in diverse settings, and populations.

Methods and analysis We will search seven databases in May 2023 with planned updates at 6 and 9 months, which will be supplemented with searches for grey literature and reference lists of included studies and contacting experts. Two reviewers will independently screen all search results for studies that evaluate one or more approaches to addressing health-related misinformation. One researcher will conduct data extraction and risk of bias assessments. which will be reviewed by a second reviewer for accuracy. We will include experimental, guasi-experimental and observational studies for any populations, settings and diseases without language or publication restrictions. We will conduct quantitative analysis if meta-analytical pooling is possible. If pooling is not possible, we will synthesise quantitative data according to outcomes and interventions addressed, and present a narrative summary of findings disaggregated by sex and/or gender, irrespective of whether differences were found.

Ethics and dissemination There are no individuals or protected health information involved and no safety issues identified. Results will be published through the Global Commission on Evidence and COVID-END websites, in a peer-reviewed journal, as well as through plain-language materials

PROSPERO registration number CRD42023421149.

INTRODUCTION

Increasing digitalisation and use of social media is a two-edged sword. It creates opportunities to rapidly communicate and disseminate information to address social challenges, and is therefore an important tool for reaching individuals and communities. 1 2 However, as

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ This living evidence synthesis will provide a comprehensive, policy relevant and evolving understanding of what is known about strategies to address health-related misinformation, and how strategies may vary by contexts and groups, including those who are more vulnerable to misinformation.
- ⇒ An interdisciplinary team of researchers and knowledge users and citizen leaders will work closely throughout the project and will use an integrated approach to knowledge translation that will ensure that findings are presented in a way that informs decisions made by government policymakers; system, organisational and professional leaders; and citizens.
- ⇒ This study will include a sex and gender perspective by documenting whether included studies comment on the inclusion or absence of sex and gender, discussing the differences in misinformation according to sex and gender, and highlighting relevant research gaps.
- ⇒ The lack of an explanatory framework might hide or underestimate some mechanisms or strategies to address misinformation, but this is a conceptual gap that we hope to fill in the future.

emphasised by the United Nations, digital technologies and social media also have the potential of introducing misinformation to citizens. Misinformation, which refers to inadvertent misleading information to which the public may be exposed and share without intent to cause harm, can delay or prevent effective care,² affect mental health,⁴ lead to misallocation health resources⁴ and/or create or exacerbate public-health crises. 45 Disinformation or malinformation are other common terms, but refer to a purposive strategy to induce false belief, channel behaviour or damage trust and can be very difficult to prove and therefore we use the term misinformation.5



Table 1 Internet sources of misinformation (table adapted from 'Public health and online misinformation: Challenges and recommendations')12

Source	Description		
Direct sources	Individuals can go straight to disreputable online domains to read information regarding health, and/or have misinformation directly disseminated to them.		
Search engines	Individuals use search engines seeking information, but can be directed to irrelevant web pages.		
User-generated content sites	 Platforms that provide an ecosystem for coproduction and consumption of content by users (eg, Yelp, Wikipedia). Social media platforms (eg, Facebook, Twitter). Visual and video platforms (eg, Instagram, Pinterest, YouTube, TikTok). Group chats (eg, WhatsApp), which are less resilient to misinformation. 		
Mobile applications	The proliferation of mobile health applications has largely been without oversight or regulation, and the quality of these applications is highly variable.		

Misinformation can affect some members of society more than others (eg, those with lower digital, numerical and health literacy and/or cognitive skills are more vulnerable to misinformation), ⁶⁷ and therefore be more exposed to health threats, leading to greater social and health inequities.⁸ A systematic review conducted in 2021 found the groups most vulnerable to health-related misinformation include younger people, those with lower educational attainment, racial minorities and social media users. The disproportionate impact on women, trans and non-binary people has also been highlighted. For instance, a report published by Plan International Australia¹⁰ shows that those groups are bombarded with stereotypes and misleading facts about their bodies and their health. This was found to lead to feeling unsafe because of online health information and questioning whether to get COVID-19 vaccines.

Extensive COVID-19-related misinformation spurred efforts to mitigate the spread of falsehoods and undermine public trust in evidence-based care. Such efforts were a focus in a report from the Broadband Commission for Sustainable Development¹¹ and in the Global Commission on Evidence.³ Individuals can engage with misinformation through different sources (see table 1).¹² In particular, while social media platforms are a key driver of misinformation, 4 13 it is not well understood since data is not publicly available for analysis and because many popular platforms (eg, Instagram, YouTube, TikTok and Pinterest), use visual content instead of texts. 12 For example, a study that analysed 800 vaccine-related Pinterest posts found that 74% were antivaccine in sentiment. 14 15

Although vaccines were the most common topic of misinformation before COVID-19, 16 other common topics for misinformation include reproductive health, substance use or smoking, non-communicable diseases, pandemics, eating disorders and medical treatments.¹⁶ Governments have employed a variety of strategies designed to debunk misinformation, including monitoring and fact-checking, economic incentives and legislative policies. 17-20 These strategies must be assessed and compared in terms of impacts and effects on health

In Public health and online misinformation: Challenges and line domains to read information regarding health, and/or have station, but can be directed to irrelevant web pages. Opportunition and consumption of content by users (eg, Yelp, Itter).

Protected by copyright, including for uses registed to misinformation.

It has largely been without oversight or regulation, and the coutcomes and behaviour change. For instance, one older evidence synthesis found that correcting misinformation has a moderate influence on belief in misinformation, rebuttals are more effective than forewarnings and appeals to coherence are more effective than fact-checking and appeals to coherence are more effective than fact-checking and appeals to coherence are more effective than fact-checking and appeals to coherence are more effective than fact-checking and appeals to coherence are more offered to be act and data mining. All trainings and innovation of misinformation challenges in particular, the rapid expansion of artificial intelligence (AI) has many implications for misinformation, which will need to be better understood. This included minings and credibility labelling). Given this, there is a need for new migh-quality and routinely updated evidence syntheses from trusted sources that assess the comparative impact of different strategies.

In 2020, the International Telecommunication Union and UNESCO, sponsored the Broadband Commission developed a report about countering digital misinformation while respecting freedom of expression. The report provides a framework explaining the stages in the misinformation process (table 2). In addition, it provides a framework for 10 potentially effective responses to misinformation and the possible intersections with freedom-of-expression rights, which is outlined in the misinformation in different settings and across diverse populations.

OBJECTIVE

OBJECTIVE

**Our objective is to synthesise and continually update empirical evidence on strategies to address health-related misinformat

Table 2 Five stages of the misinformation cycle (table adapted from 'Broadband Commission research report on 'Freedom of Expression Addressing Disinformation on the Internet')¹¹

Stage	Description of questions that arise
Instigators and beneficiaries	Motivation and goals.
Agents	Techniques (eg, bots, fake accounts or false identities).
Messages	 Formats, with three of the common ones including: Emotive claims and narratives, which often mix emotional language, lies or incomplete information, personal opinions and elements of truth. Fabricated, de-contextualised or fraudulently altered images, videos or audio. Fabricated websites and polluted data sets.
Intermediaries	Platforms (eg, dark web, social media, messaging and news media) and the platform features that are being exploited (eg, algorithms and business models).
Targets and interpreters	Who is affected (eg, citizens, scientists, politicians and journalists; organisations such as research centres and news agencies; communities such as black communities and Indigenous peoples; and systems such as electoral processes) and how they react (eg, ignoring or sharing to debunk the misinformation).

Table 3 Potential responses to misinformation (table adapted from: 'Broadband Commission research report on 'Freedom of Expression Addressing Disinformation on the Internet' and 'Governmental actions to address COVID-19 misinformation')^{11 23}

Response/ strategy	Description	Purpose of the strategy	Intersections with freedom-of- expression rights
Monitoring and fact-checking	Ongoing monitoring and timely exposing misinformation (eg, debunked claims) and fact-checking new claims. Judgement of trained professionals employed by independent organisations, even when helped by automation.	Mitigating dissemination of disinformation, false information and misinformation.	Can mitigate the risk of infringing on freedom-of-expression rights. Automation of appeal processes can infringe on freedom-of-expression rights.
Counter- misinformation campaigns	Specialised units to develop counter-narratives to challenge misinformation and mobilising online communities to spread high-quality evidence.		
Credibility labelling	Content-verification tools, web-content indicators, signposting to credible evidence sources and website-credibility labelling.	Disseminating and increasing access to accurate information. Restricting access to inaccurate information.	
Educational	Develop citizens' media/information literacy for critical- thinking and digital-verification, and journalists' information literacy.		
Curatorial	Point users to credible evidence sources, which can be used by news media, social media, messaging and search platforms.		
Narrative	Public condemnations of misinformation and recommendations to address it, often by political and societal leaders.		
Technical and algorithmic	Ranges from human learning to machine learning and other artificial intelligence approaches to identify misinformation, provide additional context and limit spread.		
Economic	Advertising bans, demonetising specific content (eg, for COVID-19) and approaches to remove misinformation incentives.	Addressing commercial fraud.	Can be misused as a form of private censorship.
Legislative and other policy	Criminalise acts of misinformation, directing internet communication companies to take down content and providing material support for credible information sources.	Criminalising expressions of disinformation.	Can be misused to weaken legitimate journalism and infringe on freedom-of-expression rights.
Investigative	Examine instigators, degree and means of spread, money involved and affected communities.		Can inform legislative and other responses.

METHODS

This living evidence synthesis will be reported according to guidance from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.²⁴

Registration

We used the PRISMA-Protocols checklist when writing our report,²⁵ and our living evidence synthesis protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO).

Literature search

We will conduct our first round of searches in May 2023. This will include seven electronic databases (MEDLINE, EMBASE, CINAHL, PsycINFO, COVID-END inventory of best evidence syntheses, Epistemonikos and medRxiv), as well as sources for grey literature (Google Scholar, Open Science Framework and greynet.org). Search terms were developed with the collaboration of a library scientist on our team (TN) using medical subject headings and text words related to forms of misinformation and interventions. The search strategy is included in online supplemental appendix 1 and will be updated at 6 and 9 months after the original search (and potentially further depending on funding availability).

Patient and public involvement

We engaged citizen partners in the development of the original grant proposal to the Canadian Institutes of Health Research, in team meetings and will continue to do so throughout the entire project, including presentation of results, interpretation and preparation of userfriendly materials for dissemination. One of our citizen partners (MS), has championed efforts for citizen engagement in evidence syntheses through COVID-END, Global Commission on Evidence and the Cochrane Consumer Network, and will lead the citizen team that includes three additional citizen partners (MB in Manitoba, CL in Quebec and JP in Nova Scotia).

Study period for the systematic review

The study commenced in May 2023 and completion is anticipated in May 2024.

Eligibility criteria

We will include original articles without language restrictions that evaluate one or more of the potential responses to health-related misinformation listed in table 3 (eg, monitoring and fact-checking, counter-misinformation campaigns, credibility labelling, educational, curational, narrative, technical and algorithmic, economic, investigative, legislative and other policy) for addressing health misinformation. We will include experimental, quasi-experimental, observational studies for any populations, settings and diseases (ie, we will not limit to only COVID-19 misinformation). Outcomes considered will include change in attitudes/behaviour, health benefits, harms and costs. We will exclude evidence syntheses but

will review their references to identify additional studies to include.

Eligibility assessment and data extraction

Retrieved studies from all electronic databases will be imported into Covidence. Two reviewers will independently screen all search results for inclusion. We will assess the titles and abstracts of the references to classify them as 'potentially relevant' or 'exclude', with disagreement resolved by consensus. We will retrieve the full text of all potentially relevant articles, and two researchers will independently review them to make a final inclusion assessment. Any disagreement at this stage will be resolved by consensus.

One reviewer will extract findings from each included ? study using a standardised form, which will be checked by a second reviewer for accuracy and consistency with the protocol. We will use a pilot exercise before extracting data where two reviewers will independently extract data from a sample of five studies to identify any areas of the form that are unclear and may require revision. The extraction form includes bibliographic information (eg, title, authors, year of publication); study design information; disease focus (COVID-19 or other focus); location of study; population addressed (eg, age of participants, users of different social media or other information platforms, gender and sex of participants and any equity consideration); type of misinformation addressed based on those outlined in table 1 (including whether it is inadvertent, purposeful or not specified); the stages of the misinformation cycle outlined in table 2; type of responses to misinformation using the framework presented in table 3; outcome(s) measured (eg, change **5** in attitudes/behaviour); and effects identified (eg, bene- ■ fits, harms, costs). We have deliberately scoped outcome measures broadly to include all potential health outcomes described in the literature.

Risk of bias in individual studies

We will appraise included studies for risk of bias and general certainty of evidence. For risk of bias, we will use the Cochrane risk of bias tool for any experimental studies. For observational study designs, we will use a version of ROBINS-I (Risk Of Bias In Non-randomised Studies - of Interventions) that was enhanced for assessment of cohort studies in a series of living evidence syntheses evaluating COVID-19 public health and social measures. We will also use the GRADE (Grading of Recommendations, Assessment, Development, and Evaluations) approach for assessing the certainty of evidence for the outcomes identified. One reviewer will independently conduct the assessments, which will be checked for accuracy by another reviewer. Any discrepancies between reviewers will be solved through consensus.

Data synthesis

Data analysis will involve quantitative analysis if metaanalytical pooling is possible, which will be based on heterogeneity (assessed with I^2) being $\leq 50\%$. We anticipate that individual studies will vary in effects observed and will estimate the pooled mean effect size using the random effects model. We will use two indicators of publication bias (Rosenthal's Fail-Safe N and Duval and Tweedie's trim and fill procedure).²⁸ For interpretation and imprecision assessment, we will use a minimally contextualised approach that considers whether the 95% CI includes the null effect, or when the point estimate is close to the null effect. We will present subgroup estimates of effect and certainty of the evidence if we observe significant differences between subgroups. For the first version of the living evidence synthesis our a priori approach will be to analyse COVID-19 and non-COVID-19 oriented strategies with further subgroup analyses determined in collaboration with our team, which includes partners from government agencies and citizens. If meta-analytical pooling is not possible, we will synthesise quantitative data according to outcomes and interventions addressed and present a narrative summary of findings using tables and charts, and disaggregating by sex and/or gender, irrespective of whether differences were found.

Updating

We anticipate producing the first version of the living evidence synthesis by the end of August 2023 with updates produced in November 2023 and March 2024. The search strategy will be adapted to include new knowledge about strategies, outcomes and sources of information to ensure we capture all the relevant scientific evidence regarding misinformation. Adopting a living approach for this evidence synthesis given that the context of misinformation is constantly shifting (eg, with increased use of AI) and may change applicability of findings over time and evidence is being produced rapidly in this area which has the potential to change conclusions about approaches over time. ²²

ETHICS AND DISSEMINATION

There are no individuals or protected health information involved and no safety issues identified.

The systematic review does not involve individuals or protected health information and there are no safety

issues identified. Each version of the living evidence synthesis will be published through the Global Commission on Evidence and COVID-END websites to ensure the most recent update is always publicly available. In addition, we publish a full version in a peer-reviewed journal that will also point to where the future updates will be housed. Moreover, we will publish plain-language materials (a summary and videos) that adopt user-friendly and engaging formats and that will be produced with our citizen partners.

Twitter Michael Wilson @Wilson_MichaelG

Collaborators The misinformation living evidence synthesis team includes: Mpho Begin, Citizen partner, Manitoba, Canada Jamie Brehaut, Ottawa Hospital Research

Institute, Canada Timothy Caulfield, University of Alberta, Canada Heather Devine, Public Health Agency of Canada Graham Dickson, LEADS Global Kelly Grimes, Canadian Health Leadership Network (CHLNet), Canada M. Mustafa Hirji, Public Health, Niagara Region, Canada Alfonso Iorio, Department of Health Research Methods, Evidence and Impact, McMaster University Nina Jetha, Public Health Agency of Canada Jennifer Kitts, Canadian Medical Association Cynthia Lisée, Citizen partner, Quebec, Canada Tamara Navarro, Health Information Research Institute, McMaster University Wendy Nicklin, Queensway Carleton Hospital, Canada Justin Presseau, Ottawa Hospital Research Institute, Canada Thomas Piggott, Peterborough Public Health, Canada Jude Porter, Citizen partner, Nova Scotia, Canada Gabrielle Plamondon, Public Health Agency of Canada Maureen Smith, Cochrane Consumer Network Bill Tholl, CHLNet, Canada.

Contributors MW is the scientific lead for the living evidence synthesis. MW, MV and JL originally conceived of the study, drafted the grant proposal and prepared this manuscript. The misinformation living evidence synthesis team includes the project co-investigators and collaborators who reviewed and provided feedback on the original proposal and this manuscript. These team members will be involved in the conduct of the living evidence synthesis through regular team meetings.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID ID

Michael Wilson http://orcid.org/0000-0002-5383-2471

REFERENCES

- 1 Grimes K, Matlow A, Tholl B, et al. Leaders supporting leaders: leaders' role in building resilience and psychologically healthy workplaces during the pandemic and beyond. Healthc Manage Forum 2022:35:213–7.
- 2 Janmohamed K, Walter N, Nyhan K, et al. Interventions to mitigate COVID-19 misinformation: A systematic review and meta-analysis. J Health Commun 2021;26:846–57.
- 3 Global Commission on Evidence to Address Societal Challenges. The Evidence Commission report: A wake-up call and path forward for decision- makers, evidence intermediaries, and impact-oriented evidence producers. Hamilton, Ontario: McMaster Health Forum, 2022
- 4 do Nascimento I, Pizarro A, Almeida J, et al. Infodemics and health misinformation: A systematic review of reviews. 2022.
- 5 Calo R, Coward C, Spiro ES, et al. How do you solve a problem like misinformation Sci Adv 2021;7:eabn0481.
- 6 Alan Turing Institute. Understanding vulnerability to online health misinformation. *Brief* 2021.
- 7 Manganello J, Gerstner G, Pergolino K, et al. The relationship of health literacy with use of Digital technology for health information: implications for public health practice. J Public Health Manag Pract 2017;23:380–7.

- 8 Jaiswal J, LoSchiavo C, Perlman DC. Misinformation and inequalitydriven mistrust in the time of COVID-19: lessons unlearned from AIDS Denialism. AIDS Behav 2020;24:2776–80.
- 9 Nan X, Wang Y, Thier K. Why do people believe health misinformation and who is at risk? A systematic review of individual differences in susceptibility to health misinformation. Soc Sci Med 2022;314:115398.
- 10 Plan international. The truth gap: the state of the world's girls. 2021. Available: https://plan-international.org/uploads/2022/02/sotwgr2021-commsreport-en.pdf
- 11 Broadband Commission research report on 'freedom of expression addressing Disinformation on the Internet'. balancing act: countering Digital Disinformation while respecting freedom of expression. 2020.
- 12 Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. *Annu Rev Public Health* 2020;41:433–51.
- 13 Broniatowski DA, Jamison AM, Qi S, et al. Weaponized health communication: Twitter Bots and Russian trolls amplify the vaccine debate. Am J Public Health 2018;108:1378–84.
- 14 Guidry JPD, Carlyle K, Messner M, et al. On pins and needles: how vaccines are portrayed on Pinterest. Vaccine 2015;33:5051–6.
- 15 Vosoughi S, Roy D, Aral S. The spread of true and false news online. science. Science 2018;359:1146–51.
- 16 Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021;23:e17187.
- 17 Au CH, Ho KKW, Chiu DKW. Stopping Healthcare misinformation: the effect of financial incentives and legislation. *Health Policy* 2021;125:627–33.
- 18 Ceretti E, Covolo L, Cappellini F, et al. Evaluating the effectiveness of Internet-based communication for public health. J Med Internet Res 2022;24:e38541.

- 19 Chou W-Y, Gaysynsky A, Vanderpool RC. The COVID-19 Misinfodemic: moving beyond fact-checking. *Health Educ Behav* 2021;48:1090198120980675:9–13.:.
- 20 Czerniak K, Pillai R, Parmar A, et al. A Scoping review of Digital health interventions for combating COVID-19 misinformation and Disinformation. J Am Med Inform Assoc 2023;30:752–60.
- 21 Walter N, Murphy ST. How to Unring the bell: A meta-analytic approach to correction of misinformation. Communication Monographs 2018;85:423–41.
- 22 Akhtar P, Ghouri AM, Khan HUR, et al. Detecting fake news and Disinformation using artificial intelligence and machine learning to avoid supply chain disruptions. Ann Oper Res 2022;2022:1–25.
- 23 Pomeranz JL, Schwid AR. Governmental actions to address COVID-19 misinformation. J Public Health Policy 2021;42:201–10.
- 24 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
- 25 Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015:4:1.
- 26 Linkins L. Critical appraisal process for assessment of public health measures for COVID-19 cohort studies. Hamilton, Canada: Health Information Research Unit, McMaster University, 2023.
- 27 McMaster Health Forum. Suite of living evidence syntheses about COVID-19 public health and social measures, . 2023Available: https:// www.mcmasterforum.org/spark-action/suite-of-living-evidencesyntheses-about-covid-19-public-health-and-social-measures
- 28 Borenstein M, Hedges LV, Higgins JPT, et al. Introduction to Meta-Analysis. In: Introduction to meta-analysis. Chichester: John Wiley & Sons, Inc, 13 March 2009.