Centre for International Governance Innovation



Digital Policy Hub - Working Paper

# Exploring the Future of Generative AI in Public Health Practice



Fall 2024 cohort

#### About the Hub

The Digital Policy Hub at CIGI is a collaborative space for emerging scholars and innovative thinkers from the social, natural and applied sciences. It provides opportunities for undergraduate and graduate students and post-doctoral and visiting fellows to share and develop research on the rapid evolution and governance of transformative technologies. The Hub is founded on transdisciplinary approaches that seek to increase understanding of the socio-economic and technological impacts of digitalization and improve the quality and relevance of related research. Core research areas include data, economy and society; artificial intelligence; outer space; digitalization, security and democracy; and the environment and natural resources.

The Digital Policy Hub working papers are the product of research related to the Hub's identified themes prepared by participants during their fellowship.

#### **Partners**

Thank you to Mitacs for its partnership and support of Digital Policy Hub fellows through the Accelerate program. We would also like to acknowledge the many universities, governments and private sector partners for their involvement allowing CIGI to offer this holistic research environment.



#### About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

Copyright © 2025 by Melissa MacKay

The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the Centre for International Governance Innovation or its Board of Directors.

Centre for International Governance Innovation and CIGI are registered trademarks.

67 Erb Street West Waterloo, ON, Canada N2L 6C2 www.cigionline.org

#### **Key Points**

- Public health plays an important role in the broader health system by promoting and protecting citizens' health, often under challenging circumstances, including declining trust, limited resources and complex information ecosystems.
- Generative artificial intelligence (AI) offers potential opportunities for public health to tailor health information to diverse populations at scale, a current challenge of public health programs and service delivery.
- Clear governance frameworks and transparent organizational public health policies prioritizing ethical principles, such as protecting autonomy, promoting trust and ensuring benevolence, are required because of critical gaps in the governance of generative AI.
- Effective use of generative AI in public health requires training, meaningful human oversight and collaboration across disciplines and diverse communities with adequate resourcing to support this in practice.
- Sustainable practices should balance environmental impacts with the technology's potential for improving health outcomes.

### Introduction

This working paper explores the application of generative AI to public health research and practice. The aim of this research was to conduct a rapid review exploring what is known about the trustworthy and responsible use of generative AI in contexts relevant to public health. This paper provides an overview of public health, AI, uses of generative AI in public health, risks and governance of AI in Canada, and the results of the rapid review, including guiding principles that can be explored as the basis for responsible and trustworthy public health organizational generative AI policies.

Public health has a critically important role within the broader health system; it refers to organizations and agencies promoting and protecting health through policy, programs and services. The public health system protects and promotes health within the context of complex health issues, challenging information and digital ecosystems, public distrust, and limited resources. Trust in science and public health has been declining in recent years (Caulfield et al. 2021; Kennedy, Tyson and Funk 2022), with issues such as poor communication during the COVID-19 pandemic further eroding trust in public health and government officials (Siegrist and Zingg 2014; MacKay et al. 2022).

The accuracy and credibility of sources used for health information play a significant role in health outcomes. Online sources — the primary health information resource for many individuals — are abundant but often unreliable. For instance, there are 70,000 healthrelated Google searches per minute worldwide (Murphy 2019), and many people turn to these sources before consulting health-care providers (Swoboda et al. 2018).

Although online health information is widely available, false and misleading information (hereafter referred to as mis- and disinformation) can lead to increased worry and poor health decision making (Korshakova, Marsh and Kleinberg 2022). Although mis- and disinformation occurs in all communication channels, the rapid diffusion of information

through digital channels facilitates its proliferation (Burke-Garcia and Soskin Hicks 2024). As a result, people often make health decisions based on inaccurate, false or irrelevant information that is not intended for their particular needs or circumstances.

AI tools offer promising opportunities to enhance public health outcomes. AI has demonstrated potential in precision medicine and can similarly help public health tailor activities to specific populations and needs (Fisher and Rosella 2022). AI models can support efficient health communication, predictive analysis and optimization frameworks, which may be particularly useful for health authorities with limited resources (ibid.). Limited resources have contributed to a significant training gap in the public health workforce, particularly in using current and emerging technologies, as well as difficulties with the delivery of high-quality programs and services (Canadian Public Health Association 2022). However, AI systems must be carefully designed and deployed, as biases related to race, gender, sexual orientation, class and other identities are being integrated into their outputs (King 2022; Akselrod 2021).

Generative AI is experiencing rapid development and use due to its transformative potential. It can generate text, images, audio, videos and computer code responding to user prompts (Office of the Privacy Commissioner of Canada 2023). This capability has captured public attention for its ability to produce tailored, evidence-based health information at scale and speed (Burke-Garcia and Soskin Hicks 2024).

However, generative AI is not without risks, particularly in health contexts. While it can generate accurate and tailored information, generative AI is also prone to errors, including fabricating sources, providing unsafe advice and introducing factual inaccuracies.<sup>1</sup> Generative AI models can also reinforce and exacerbate societal biases without safeguards in place (Nicoletti and Bass 2023). These issues exacerbate the challenges posed by mis- and disinformation.

Public health should explore the opportunity to leverage generative AI for good, using it to support and promote people's health and well-being at the population level whenever possible. Generative AI may be able to help public health meet the widespread need for high-quality, empathic, tailored health information. By producing accurate, culturally sensitive, tailored and timely communication, generative AI could help counteract the negative impacts of mis- and disinformation and improve the relevance of health information. These opportunities are explored in more detail below, including tailored communication, content analyses, thematic analyses, evidence syntheses and chatbots.

The stakes are too high for public health to ignore the potential of generative AI, particularly in the context of communication. This research uses a rapid review to explore how generative AI can be ethically and responsibly integrated into Canadian public health research and practice. Key considerations that will be explored include:

- **Equity:** Ensuring AI-generated content is inclusive and addresses the needs of diverse populations, including and especially those that have been equity denied.
- Accuracy: Public health must ensure that the generated content is true and correct.

<sup>1</sup> See www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guideuse-generative-ai.html.

- **Transparency:** Making the role of AI in generating health information clear to the public.
- Accountability: Hold developers and implementers responsible for the ethical use of generative AI.
- **Collaboration:** Public health should collaborate with developers and other sectors to develop and improve generative AI models.
- **Human-centredness:** Utilizing generative AI as a complementary tool where public health researchers and practitioners actively participate in key steps rather than being replaced by the models.

By observing these considerations, public health can more safely consider how to harness the potential of generative AI to improve communication, build trust and ultimately enhance health outcomes across Canada. These guiding principles are further discussed in the results section of this working paper.

# Faster, Tailored and Scalable Approaches to Public Health Communication

There are many tools at our disposal to enact tailored and more efficient and effective public health communication at scale. Tailored health information is differentiated based on factors such as reading level, language, worldviews, behaviours and identities. The tools include:

- **Tailored communication:** Information relevant to individuals and communities based on their values, needs, wants, behaviours and other contextual factors.
- **Speed and scale:** Models provide the opportunity to generate various types of communication products at a speed and scale that cannot be replicated by human practitioners.
- **Content analyses:** Identify, assess and make sense of large amounts and varying types of data and data sources.
- **Thematic analyses:** Identify and interpret patterns of meaning or themes within data.
- **Evidence synthesis:** Bring together information from various sources on the same topic to inform decisions.
- **Chatbots:** Provide instant, personalized, 24/7 information support that mimics human interactions.

Use cases, or applications of generative AI models to practice, are fairly prevalent in health-care-related research, but research specific to using generative AI in public

health is in its infancy and quite limited. Below, use cases relevant to public health communication in the literature are explored.

#### **Tailored Health Information**

Health information is essential for decisions that impact overall health and wellbeing. The quality and accuracy of this information make a difference in health equity and health-related outcomes. Generative AI provides the opportunity to tailor health information to the language, literacy level, preferences and other considerations of individuals and communities. Research specific to public health is needed in this area, especially because of ethical issues, which will be discussed later.

One study found that generative AI can produce summaries of scientific articles that use clear language (short words and short sentences) and are more accessible to the general public compared to human-generated summaries (Markowitz 2024). Importantly, the AI-generated summaries were also perceived as more credible and trustworthy (ibid.). Participants were able to summarize the science more accurately after reading generative AI-produced summaries as opposed to human-generated summaries of the same science (ibid.).

Another body of research is focused on the ability of generative AI to be an opinion leader. Opinion leaders are powerful for health information dissemination that is credible and tailored and that can drive health behaviours (Burke-Garcia and Soskin Hicks 2024). Opinion leaders are known to communities, considered trustworthy and knowledgeable, and can engage in empathic communication (ibid.). Researchers are developing what they call "health communication AI" that leverages what is known about behaviour, opinion leadership and current technology to provide accurate and trustworthy health information to diverse communities (ibid.).

Generative AI may be useful in creating first-draft versions of evidence-based health messages for diverse audiences. Models such as ChatGPT or Copilot can be prompted to incorporate relevant models and frameworks, message frame types (loss and gain frames, for example) and produce various audience segments and different formats. Practitioners must always carefully review and edit the outputs, especially for referencing and ethical issues.

#### **Content Analysis**

Content analysis can be used to analyze data from interviews, focus groups, openended survey responses and other formats. It allows for a systematic approach to understanding how much the data corresponds to a theory or framework or how many times the data mentions different topics of interest. This systematic approach to analyzing data provides opportunities for technology to be of assistance. Generative AI may increase the efficiency of qualitative research, including content analysis, but there is still a need for effective prompts, understanding the responses and ensuring the trustworthiness of the analysis.

One study examined ChatGPT's ability to conduct a qualitative content analysis of online forum data on reducing sugar consumption. Researchers developed the prompts and prepared the data, while ChatGPT conducted the content analysis. ChatGPT was

able to identify instances of behaviour changes related to discussion of reducing sugar consumption across the data (Bijker et al. 2024). The results of the ChatGPT content analysis were compared to human analysis and found to be fairly reliable (ibid.).

Another study used a range of data sources, including social media, news and technical reports, and conducted a generative AI-assisted content analysis. ChatGPT was prompted to look for specific topics within the data, such as the presence of mis- and disinformation and references to ideas, people and places. This study found that ChatGPT is just as accurate as humans in identifying the topics, and the model completed the task 36 percent faster (Chew et al. 2023).

In short, generative AI shows promise for content analysis, though not as a replacement for humans but as a tool to support their methodology. It can combine different types of data and complete a content analysis for the presence of themes, mis- and disinformation and even the mention of opinion leaders just as accurately and a great deal faster than humans. It can be used by public health researchers and practitioners to understand health information needs and wants as well as investigate what type of misand disinformation is circulating.

#### **Thematic Analysis**

Like content analysis, thematic analysis can analyze data from interviews, focus groups and open-ended survey responses. However, it goes beyond counting words or phrases to understanding the underlying meaning in the data, including concepts, experiences and ideas.

One study used ChatGPT to conduct a thematic analysis based on health-related information. It assessed the ability of the model to first identify patterns in the data, then extract themes and subsequent quotations to represent each theme (Lee et al. 2024). Although the researchers found that ChatGPT showed the potential for conducting a thematic analysis, human oversight and input are still required (ibid.).

As with content analysis, generative AI can be a valuable and fast tool to help members of the analysis team, contributing to comprehension and overall efficiency. However, it also requires human oversight and should complement the work of practitioners rather than replace them. Thematic analysis can help identify community values and information needs, as well as experiences related to health.

#### **Generative AI Chatbots**

Chatbots mimic human conversation and interaction through written, oral and visual communication. They enable accessible, instant and engaging health information services, although challenges related to sustainability, inflexibility and adherence have been reported (Aggarwal et al. 2023). A review found high efficacy related to chatbots' ability to promote healthy lifestyles, smoking cessation and reduction in substance misuse, which all have high relevance to public health (ibid.).

The World Health Organization (WHO) launched a digital health promoter, which is a chatbot powered by generative AI. "SARAH" can be accessed 24 hours a day in eight different languages and provides information across different public health topics such as healthy behaviours and mental health. It leverages generative AI so that it can engage in two-way conversations and provide nuanced, tailored and empathic responses (WHO

2024b). It is meant to provide advice on some of the biggest causes of death worldwide including cancer, heart disease and diabetes (ibid.). However, a recent article reports that its outputs are very inconsistent and often provide poor quality answers and broken links (O'Neill and Eccles 2024).

Other chatbots have been found to deliver quality and empathic responses in health care. One study found that participants preferred chatbot responses to those of physicians across the vast majority of responses (Ayers et al. 2023). Chatbot responses were rated as higher quality and significantly more empathetic than physician-generated responses (ibid.).

This technology promises two-way communication and instant, tailored and accurate information, but also requires additional research and development.

# **Risks Associated with AI**

Despite the promising use cases and potential for more efficient and effective public health communication research and practice, a number of risks have been welldocumented with respect to AI systems in general. According to Mark R. Miller, Connie Moon Sehat and Robert Jennings (2024), these risk include:

- **Privacy concerns:** AI models can collect and analyze personal health data, which can be de-anonymized and is vulnerable to security breaches.
- **Bias:** AI models can exacerbate bias and discrimination based on the data used to train them and the developers who build them.
- Accuracy: AI models may include inaccurate, outdated or made-up information (hallucinations) and references.
- **Mis- and disinformation:** AI models can produce myths, conspiracy theories and inaccurate and misleading information. Examples include vaccine mis- and disinformation, false cures and treatments for health issues, supposed health risks of 5G technology, diet myths, perpetuation of the stigma around mental health and false information about chronic diseases.
- Workforce impacts: Impacts on the health related to the workforce include possible job losses, the impact on existing jobs due to additional training and task adjustment and an overdependence on generative AI.
- **Environmental impacts:** AI requires far more power than traditional uses of the internet, such as Google searches.
- **Technical challenges:** AI systems require ongoing competencies and resources for use and maintenance.

# Governance of AI in Public Health in Canada

Integrating AI, including generative AI, into public health in Canada requires robust governance to ensure compliance with privacy laws, ethical standards and public trust. Public health organizations using generative AI must align their activities with applicable privacy legislation, such as the Personal Information Protection and Electronic Documents Act (PIPEDA), and health-specific regulations to protect sensitive health information (Office of the Privacy Commissioner of Canada 2023). Beyond legal compliance, organizations must adopt frameworks that emphasize responsibility and trustworthiness to safeguard human rights and dignity while ensuring the outputs and impacts are beneficial (ibid.).

# Challenges to Responsible and Trustworthy AI Use in Health

Despite its potential, trust in AI use, including generative AI, in health-care settings is low (Gliadkovskaya 2023). Research often focuses on the demographic characteristics of users and the performance characteristics of AI systems to understand trust in these technologies (Steerling et al. 2023). However, this narrow focus overlooks critical contextual factors, values and implementation considerations such as the intended use of generative AI. A holistic perspective is essential to build a nuanced understanding of trust and trustworthiness in generative AI applications (ibid.).

#### **Regulatory Landscape for AI**

Privacy legislation in Canada varies depending on the nature of the organization (public or private) and the purpose of AI use (Office of the Privacy Commissioner of Canada 2023). While Canada lacks a regulatory framework specific to AI, the proposed Artificial Intelligence and Data Act (AIDA) aims to address this gap. AIDA seeks to ensure that AI systems in Canada are safe and non-discriminatory and that they are used responsibly, aiming to hold organizations accountable for misuse.<sup>2</sup>

However, as the act currently stands, it has important limitations. According to Kristen Thomasen (2023), these are:

- Harm definition: Harms are primarily defined based on individual human rights and do not adequately address population- and community-level concerns or indirect harms inherent to generative AI outputs (Thomasen 2023). For example, algorithms using postal codes to infer criminal activity risk or hiring systems trained on male-centric data sets illustrate the potential for systemic bias and inequity (ibid.).
- **Government use:** Government uses of AI are currently excluded from the act, and the focus is on commercialization rather than equity issues (ibid.).
- **Emphasis on commercialization:** AIDA prioritizes the commercialization of AI systems over equity considerations, leaving critical gaps in addressing fairness and inclusion in public health contexts (ibid.).

2 See https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act.

#### **Guidance on Generative AI Use**

#### Government of Canada Guide on the Use of Generative AI

The Government of Canada published a guide on the use of generative artificial intelligence in 2024. The guide advises federal institutions to explore generative AI tools cautiously, emphasizing risk evaluation and effective management before implementation. Low-risk uses include using generative AI to draft an email or edit documents that will go through further human revisions.<sup>3</sup> Higher risk uses include using chatbots or incorporating client information into a summary.<sup>4</sup>

In the responsible and trustworthy use of generative AI, organizations must prioritize fairness by avoiding biases and engaging stakeholders, ensuring accountability through oversight and ethical content generation, and supporting security by safeguarding privacy and managing cybersecurity risks.<sup>5</sup> Transparency requires the clear identification of AI-generated content and accessible explanations of its use.<sup>6</sup> Practitioners must be educated on generative AI's strengths and limitations while ensuring the tools are relevant to user and organizational needs.<sup>7</sup>

#### **Canadian Centre for Cyber Security Generative AI**

The Canadian Centre for Cyber Security (2023) also outlines principles for generating quality and trustworthy content while mitigating concerns. They include:

- **Establishing generative AI use policies:** Organizations should develop clear policies on the types of content that can be generated and guidelines for using that technology. These policies should include oversight and review processes to ensure appropriate and ethical use.
- **Training data sets:** Data sets should be diverse, representative of communities and reviewed by a diverse team to mitigate potential biases.
- Using best practices for tools and vendors: Only generative AI tools from vendors with strong security practices and transparency in their development should be used.
- **Mitigating privacy concerns:** Avoid using personally identifiable information or private organizational information in prompts unless the tool maintains data privacy.

## WHO Ethics and Governance of AI for Health: Guidance on Large Language Models

The WHO recently extended its guidance on AI to include large language models, a subset of generative AI models. The WHO consensus on ethical principles for AI use in health include protecting autonomy; promoting human well-being, human safety and public interest; ensuring transparency, explainability and intelligibility; fostering responsibility and accountability; ensuring inclusiveness and equity; and promoting

- 6 Ibid.
- 7 Ibid.

<sup>3</sup> See www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guideuse-generative-ai.html.

<sup>4</sup> Ibid.

<sup>5</sup> Ibid.

AI that is responsive and sustainable (WHO 2024a). While much of the guidance relates to health care and medicine, these principles are also relevant to public health. International governance of AI is recommended where governments collectively develop international rules for AI (ibid.).

## Strengthening AI Governance: Addressing Gaps and Advancing Equitable, Adaptive Policies

AI policies and regulations at the national and international levels provide essential frameworks but also reveal critical gaps that must be addressed. For example, PIPEDA offers strong protections for personal data but lacks provisions tailored to AI-specific challenges, such as algorithmic bias or systemic harms.<sup>8</sup> Similarly, AIDA is a promising step toward responsible AI governance, yet it prioritizes commercialization and individual harms over equity and public sector applications, excluding government AI use entirely (Kamyabi et al. 2024). The federal Directive on Automated Decision-Making ensures fairness and transparency for certain public sector AI systems but is limited in scope and enforcement (Attard-Frost, Brandusescu and Lyons 2024). The Canadian Centre for Cyber Security effectively addresses cybersecurity risks but does not engage with broader ethical concerns (Canadian Centre for Cyber Security 2023). At the international level, the WHO Guidance on Large Language Models provides valuable ethical principles but lacks binding mechanisms and specificity for localized applications (WHO 2024a).

To strengthen these frameworks, policies should integrate equity and inclusivity, explicitly address systemic inequities and define clear roles and accountability mechanisms across sectors. Expanding the scope to include all public sector applications and creating mechanisms to update regulations aligned with technological advancements are also critical. These frameworks demonstrate the need for organizational policies guiding AI and generative AI use for a more comprehensive, dynamic and inclusive approach to AI governance.

#### **Need for Organizational Generative AI Policies**

Canada needs policies and guidelines for AI use in public health and processes that evaluate and assess potential bias and equity issues.<sup>9</sup> These policies must address fundamental issues such as bias, equity and transparency, ensuring that AI systems align with public health goals and values. Relying solely on legislation is insufficient; organizations must develop internal policies that reflect ethical principles and anticipate gaps in existing or proposed regulations.

Organizational policies can be adapted more rapidly than legislation, enabling public health to respond more effectively to the ongoing evolution of AI tools and applications. This flexibility allows organizations to pilot new technologies, address emerging risks and implement updates without the lengthy processes required for legislative changes. Organizational policies around generative AI can foster a culture of innovation, ensuring that public health agencies explore new and emerging tools ethically and equitably. By addressing challenges such as data privacy, bias and equity in a timely manner, organizational policies can better align with complex public health needs and technological advancements.

<sup>8</sup> See https://ised-isde.canada.ca/site/innovation-better-canada/en/canadas-digital-charter/strengthening-privacy-digital-age.

<sup>9</sup> See Fisher and Rosella (2022); https://cihr-irsc.gc.ca/e/53244.html

Policies must include robust processes to evaluate and mitigate potential biases and harms. Transparency and accountability are central. This is particularly vital in public health, where AI models can exacerbate health disparities. Public health organizations must clearly communicate how, why and when generative AI is being used<sup>10</sup> and establish mechanisms for accountability. Guiding principles for generative AI use should foster trust, contribute to positive public health outcomes and be reflected in policy. The aim of this research is to explore what is known about the responsible and trustworthy use of generative AI in contexts relevant to public health within the current published literature. While some international and federal guidance on generative AI use exists, no research has specifically examined the literature related to the responsible and trustworthy use of generative AI in public health. The results of this research serve to extend the current guidance with evidence-based recommendations specific to generative AI, and its applications in public health communication.

# Method

A rapid review of peer-reviewed and grey literature was implemented to understand the guiding principles associated with the responsible and trustworthy use of generative AI in public health. Rapid reviews produce results in a more timely and resource-effective manner to answer broad questions that have relevance to policy (Smela et al. 2023). Controlled vocabulary and keywords related to generative AI, trust and responsible use and public health were used to search Medline via OVID, Web of Science, PsycINFO and Compendex in October 2024. Grey literature was also searched using combinations of keywords in Google and searching up to 10 pages for each keyword combination.

10 See Office of the Privacy Commissioner of Canada (2023); https://cihr-irsc.gc.ca/e/53244.html; www.canada.ca/en/ government/system/digital-government/digital-government-innovations/responsible-use-ai.html.

#### Table 1: Inclusion and Exclusion Criteria for Rapid Review

Inclusion Criteria	Exclusion Criteria
All methods	No substantial focus on generative Al use in public health
Original peer-reviewed articles and grey literature	Commentaries, editorials, viewpoints
English language	Articles focused on clinical, medical or health-care topics with no relevance to public health
2014 or later	Articles only focused on generative AI development
Relevant to generative AI use in a public health context (e.g., relevant to population- level health, health promotion, etc.)	
Explore factors related to the responsible and trustworthy use of Al	

Source: Author.

The author completed title and abstract and full-text screening, with a second reviewer verifying 10 percent of the data set at each stage. Then, in collaboration with the second researcher, the author also completed data extraction and thematic analysis of the results.

# Results

Overall, 10 articles were included based on the inclusion criteria. Of the 10 articles, nine are journal articles and one is grey literature. The methods included two reviews, three descriptive summaries, two interviews and thematic analyses, one survey and one cross-sectional analysis. Key themes and related data are explored below and relate to the key considerations and guiding principles for the responsible and trustworthy use of generative AI in public health. The research is very limited at this time, and these themes represent a starting point for public health organizations to explore generative AI use and policies.

#### Key Themes in the Data

#### Applications of AI in Public Health

• Generative AI supports public health by analyzing data sets for epidemiological purposes, such as tracking virus spread, predicting outbreaks and informing strategies during crises such as the COVID-19 pandemic (Alhur 2024).

• Generative AI delivers tailored, empathetic health communication at scale, addressing misinformation and promoting public health awareness (Burke-Garcia and Soskin Hicks 2024; Sallam et al. 2023).

#### **Trust and Transparency**

- Trust is central to generative AI adoption in public health. Trustworthy generative AI systems require transparency about their use, accuracy and limitations (Chen, Kuo and Chang 2024; Hussain, Wang and Li 2024).
- Opinion leaders, such as social media influencers, serve as models for fostering trust in health communication AI by establishing credibility and emotional connection (Burke-Garcia and Soskin Hicks 2024).
- Generative AI has the ability to reinforce trust by citing credible health organizations and refuting mis- and disinformation (Sallam et al. 2023).

#### **Ethical and Regulatory Challenges**

- Ethical concerns include bias, equity, privacy and accountability. AI systems must protect autonomy, avoid harm and adhere to regulatory standards (WHO 2024a; Sallam et al. 2023).
- Reliance on large internet-based data sets poses risks associated with biased and incomplete training data (ibid.).
- Current regulations and guidelines are fragmented, underscoring the need for clearer governance frameworks tailored to AI in health care and public health (Ning et al. 2024).

#### **Bias and Equity Issues**

- Data quality and representation are critical to avoiding algorithmic bias that perpetuates health disparities (Alhur 2024).
- Outputs must be culturally respectful, accurate and unbiased to avoid harm and maintain trust between public health and diverse communities (Burke-Garcia and Soskin Hicks 2024).

#### **Misinformation and Safeguards**

- AI can both propagate and combat misinformation. Effective safeguards, such as robust validation and monitoring, are necessary to prevent the spread of health disinformation (ibid.; Chen, Kuo and Chang 2024; Sallam et al. 2023).
- Gaps in safeguards and inconsistent adherence to ethical principles highlight the need for standards and oversight in generative AI applications (Menz et al. 2024).

#### **Education and Training**

- Public health practitioners require training to effectively integrate generative AI into practice, emphasizing limitations, ethical use and trust-building (Alhur 2024; WHO 2024a; Burke-Garcia and Soskin Hicks 2024).
- Institutions must invest in innovative teaching methods to prepare practitioners to use generative AI while navigating the complexities of public health (Chen, Kuo and Chang 2024).

#### **Challenges in Implementation**

- Resource constraints, including financial, technological and human resources, limit generative AI adoption in public health (Alhur 2024; Morgenstern et al. 2021).
- Public health partnerships with technology companies are necessary to catalyze generative AI innovation and address expertise gaps (Burke-Garcia and Soskin Hicks 2024; Hussain, Wang and Li 2024).

#### **Human-Centred Generative AI Use**

- Generative AI should complement, not replace, human judgement, particularly in sensitive health-related decisions (WHO 2024a; Alhur 2024; Burke-Garcia and Soskin Hicks 2024; Hussain, Wang and Li 2024).
- Human oversight remains essential to ensure generative AI applications align with community values (Alhur 2024).

#### **Privacy and Security**

- Protecting data privacy through anonymization, encryption and adherence to regulations is critical for maintaining trust (Menz et al. 2024; Hussain, Wang and Li 2024; Alhur 2024).
- Cybersecurity risks and data misuse pose significant threats to the responsible deployment of generative AI (Alhur 2024; Hussain, Wang and Li 2024; Menz et al. 2024).

#### Sustainability and Scalability

- Generative AI solutions must be sustainable, accessible and adaptable to various public health settings, including low-resource environments (Chen, Kuo and Chang 2024; WHO 2024a; Morgenstern et al. 2021).
- Generative AI technologies should align with broader efforts to support the sustainability of health systems and the environment (WHO 2024a).
- Overdependence on generative AI systems could undermine trust and impact workforce dynamics, necessitating a balanced approach to integration (ibid.).

#### **Future Opportunities and Risks**

- Generative AI offers opportunities for personalized health promotion and nuanced communication on controversial health topics such as vaccination (Burke-Garcia and Soskin Hicks 2024).
- Risks include skills degradation (WHO 2024a; Chen, Kuo and Chang 2024) and the psychological toll on individuals involved in monitoring or reviewing generative AI outputs that may be abusive or disturbing (WHO 2024a).

## Evidence, Guidelines and Guiding Principles for the Responsible and Trustworthy Use of Generative Al

Developing organizational policies for constantly and rapidly changing technology is difficult, but generative AI provides a tool that can help public health professionals be more efficient and effective in their work, making this technology essential to address. This section synthesizes the use cases, risks, current guidance on generative AI use and evidence from the rapid review to understand what is currently known about the principles associated with the responsible and trustworthy use of generative AI in public health. Proactive, collaborative and ethical approaches to integrating generative AI in public health must focus on human oversight, transparency, accountability and sustainability to ensure equitable health benefits while mitigating risks. The research is in its infancy and this working paper provides a starting point by which public health organizations can consider their use of generative AI and the policies that can guide the responsible and trustworthy use of this technology in practice.

#### **Human Capital Considerations**

- **Meaningful human oversight:** Policies for generative AI must describe the various points of human interaction with generative AI models and outputs and outline where human decisions must be made. Human oversight ensures that generative AI use aligns with public health values and community priorities, which fosters trust and accountability. Critical thinking by practitioners is essential to the ethical and effective use of generative AI.
- Education and training: Students and practitioners require training to effectively use and monitor generative AI. Institutions and organizations must use innovative teaching methods to equip students and practitioners with the competencies needed to make informed decisions about the technology's limitations and ethical use.
- **Interdisciplinary collaboration:** Digital fluency across public health practice is necessary so that practitioners understand the benefits and limitations of generative AI. Collaboration among various practitioners and disciplines, including informatics, computer science, ethicists and others, is needed to share our collective

understanding of and feedback on this technology that can be used for further responsible and ethical development of generative AI in public health.

#### **Governance Considerations**

- Accountability: Public health organizations are responsible for the outputs from generative AI used by public health. Public health plays a central role in decisions regarding the use of generative AI and shares responsibility for addressing any harms related to its implementation. Developers, policy makers and other stakeholders must also be accountable, working collaboratively to ensure ethical and equitable outcomes.
- Ethical and regulatory challenges: Bias, equity, privacy and accountability are significant ethical considerations for generative AI use in public health. While generative AI can contribute to developing, tailoring and synthesizing health messages at scale, there are significant risks. Current regulations remain fragmented and public health organizations must develop policies to guide generative AI use. Ethical principles such as protecting autonomy and avoiding harm should guide policy development and generative AI use. Moving beyond non-malevolence (doing no harm) to benevolence (doing good) puts the focus on generative AI squarely on positively contributing to health outcomes and overcoming health inequities.
- **Privacy and security:** Protecting privacy through anonymization, encryption and regulatory compliance is essential to maintaining public trust in generative AI and doing no harm. Transparent policies can mitigate these risks while ensuring ethical use.
- **Sustainability:** Judicious use of generative AI is important to address the technology's environmental impacts. Resource constraints, including finances and human competence, are additional considerations for the sustainable use of generative AI.

#### Human-Centredness

- **Trust and transparency:** Trust is foundational to the success of public health, including the successful use of generative AI. Transparency regarding the use, accuracy and limitations of the technology is critical to maintaining trust in diverse communities. Emotionally salient, credible, accurate and tailored health messages developed by generative AI can contribute to trust and help overcome health-related mis- and disinformation.
- **Participatory use:** Considerations about the use of generative AI should be designed with diverse communities to avoid exacerbating health inequities.

# Recommendations

- Human-centredness and oversight woven throughout policies are vital for trust and the responsible use of generative AI.
- Organizational policies must reflect guardrails to address intersecting issues such as privacy, bias, duplicated content and mis- and disinformation.
- Organizational policies must explicitly incorporate ethical principles and risks and align with organizational and societal values to promote autonomy, equity, transparency and trust.
- Education and training to enhance digital fluency and ethical understanding of generative AI are needed to upskill the public health workforce.
- Sustainable generative AI practices must balance environmental impacts with the potential for increased efficiency and effectiveness of initiatives.
- Collaboration with other disciplines and communities should be undertaken to promote the understanding and responsible use of generative AI for equitable outcomes.

#### Acknowledgements

My sincere thank you to my supervisor, Jennifer McWhirter, for her continued support and mentorship throughout my postdoctoral experiences. I am grateful to Anjali Kukan, a master's of public health student, who made invaluable contributions to the rapid review, including the search protocol, screening, thematic analysis and corresponding manuscript. Thank you to my CIGI mentor, Bianca Wylie, and CIGI postdoctoral fellow Elia Rasky, for their excellent peer review.

#### About the Author

Melissa MacKay (she/her), MPH, Ph.D., is a Digital Policy Hub post-doctoral fellow who specializes in health, risk and crisis communication, as well as health promotion. Previously a postdoctoral scholar at the University of Guelph's Health by Design Lab, she focused on developing a modernized health communication framework and using artificial intelligence for tailored health communication. Her research employs qualitative and mixed methods and evidence synthesis, and emphasizes knowledge mobilization for impact. Her dissertation examined trust in public health communication during the COVID-19 pandemic and she is the lead author of the *Social Media Crisis Communication Guidebook for Public Health*. Melissa teaches in the MPH program at the University of Guelph and has extensive experience in public health and knowledge mobilization.

# **Works Cited**

- Aggarwal, Abhishek, Cheuk Chi Tam, Dezhi Wu, Xiaoming Li and Shan Qiao. 2023. "Artificial Intelligence–Based Chatbots for Promoting Health Behavioral Changes: Systematic Review." Journal of Medical Internet Research 25: e40789. https://doi.org/10.2196/40789.
- Akselrod, Olga. 2021. "How Artificial Intelligence Can Deepen Racial and Economic Inequities." American Civil Liberties Union, July 13. www.aclu.org/news/privacy-technology/ how-artificial-intelligence-can-deepen-racial-and-economic-inequities.
- Alhur, Anas. 2024. "Redefining Healthcare With Artificial Intelligence (AI): The Contributions of ChatGPT, Gemini, and Co-pilot." *Cureus* 16 (4): e57795. https://doi.org/10.7759/cureus.57795.
- Attard-Frost, Blair, Ana Brandusescu and Kelly Lyons. 2024. "The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives." *Government Information Quarterly* 41 (2): 101929. https://doi.org/10.1016/j.giq.2024.101929.
- Ayers, John W., Adam Poliak, Mark Dredze, Eric C. Leas, Zechariah Zhu, Jessica B. Kelley, Dennis J. Faix et al. 2023. "Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum." *JAMA Internal Medicine* 183 (6): 589–96. https://doi.org/10.1001/jamainternmed.2023.1838.
- Bijker, Rimke, Stephanie S. Merkouris, Nicki A. Dowling and Simone N. Rodda. 2024. "ChatGPT for Automated Qualitative Research: Content Analysis." *Journal of Medical Internet Research* 26: e59050. https://doi.org/10.2196/59050.
- Burke-Garcia, Amelia and Rebecca Soskin Hicks. 2024. "Scaling the Idea of Opinion Leadership to Address Health Misinformation: The Case for 'Health Communication AI." *Journal of Health Communication* 29 (6): 396–99. https://doi.org/10.1080/10810730.2024.2357575.
- Canadian Centre for Cyber Security. 2023. "Generative artificial intelligence (AI)." ITSAP.00.041. July. www.cyber.gc.ca/en/guidance/generative-artificial-intelligence-ai-itsap00041
- Canadian Public Health Association. 2022. Strengthening Public Health Systems in Canada. Ottawa, ON: Canadian Public Health Association. www.cpha.ca/sites/default/files/uploads/ advocacy/strengthen/strengthening-ph-systems-brief-e.pdf.
- Caulfield, Timothy, Tania Bubela, Jonathan Kimmelman and Vardit Ravitsky. 2021. "Let's do better: public representations of COVID-19 science." *FACETS* 6 (1): 403–23. https://doi.org/10.1139/facets-2021-0018.
- Chen, Su-Yen, H. Y. Kuo and Shu-Hao Chang. 2024. "Perceptions of ChatGPT in healthcare: usefulness, trust, and risk." *Frontiers in Public Health* 12: 1457131. https://doi.org/10.3389/ fpubh.2024.1457131.
- Chew, Robert, John Bollenbacher, Michael Wenger, Jessica Speer and Annice Kim. 2023. "LLM-Assisted Content Analysis: Using Large Language Models to Support Deductive Coding." Preprint, *arXiv*, June 23. https://doi.org/10.48550/arXiv.2306.14924.
- Fisher, Stacey and Laura C. Rosella. 2022. "Priorities for successful use of artificial intelligence by public health organizations: a literature review." *BioMed Central Public Health* 22, 2146. https://doi.org/10.1186/s12889-022-14422-z.
- Gliadkovskaya, Anastassia. 2023. "Most U.S. patients don't trust AI in healthcare settings, feel transparency would help, Carta survey finds." Fierce Healthcare, October 20. www.fiercehealthcare.com/health-tech/most-americans-dont-trust-ai-healthcare-setting-carta-survey-finds.

- Hussain, Talib, Dake Wang and Benqian Li. 2024. "The influence of the COVID-19 pandemic on the adoption and impact of AI ChatGPT: Challenges, applications, and ethical considerations." *Acta Psychologica* 246, 104264. https://doi.org/10.1016/ j.actpsy.2024.104264.
- Kamyabi, Alireza, Alex Choi, Kevin Lam, Tina Purnat, Juan Solorzano, Manik Saini, Geoffrey McKee et al. 2024. "Submission to the Standing Committee on Industry and Technology on Bill C-27, *The Digital Charter Implementation Act, 2022.*" Public Health Association of British Columbia. www.ourcommons.ca/Content/Committee/441/INDU/Brief/BR12921212/ br-external/PublicHealthAssociationOfBritishColumbia-e.pdf.

Kennedy, Brian, Alec Tyson, and Cary Funk. 2022. Americans' Trust in Scientists, Other Groups Declines. Pew Research Center Science & Society, February 15. www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groupsdeclines/.

- King, Morgan. 2022. "Harmful biases in artificial intelligence." *The Lancet Psychiatry* 9 (11): e48. https://doi.org/10.1016/S2215-0366(22)00312-1.
- Korshakova, Elena, Jessecae K. Marsh and Samantha Kleinberg. 2022. "Health Information Sourcing and Health Knowledge Quality: Repeated Cross-Sectional Survey." *Journal of Medical Internet Research Formative Research* 6 (9): e39274. https://doi.org/10.2196/39274.
- Lee, V. Vien, Stephanie C. C. van der Lubbe, Lay Hoon Goh and Jose Maria Valderas. 2024. "Harnessing ChatGPT for Thematic Analysis: Are We Ready?" *Journal of Medical Internet Research* 26 (1): e54974. https://doi.org/10.2196/54974.
- MacKay, Melissa, Taylor Colangeli, Jillian Jaworski, Caitlin Ford, Andrew Papadopoulos, Jennifer McWhirter and Daniel Gillis. 2022. Social Media Crisis Communication Guidebook for Public Health. Guelph, ON: University of Guelph. https://hdl.handle.net/10214/27017.
- Markowitz, David M. 2024. "From complexity to clarity: How AI enhances perceptions of scientists and the public's understanding of science." *PNAS Nexus* 3 (9): pgae387. https://doi.org/10.1093/pnasnexus/pgae387.
- Menz, Bradley D., Nicole M. Kuderer, Stephen Bacchi, Natansh D. Modi, Benjamin Chin-Yee, Tiancheng Hu, Ceara Rickard et al. 2024. "Current safeguards, risk mitigation, and transparency measures of large language models against the generation of health disinformation: repeated cross sectional analysis." *British Medical Journal* 384: e078538. https://doi.org/10.1136/bmj-2023-078538.
- Miller, Mark R., Connie Moon Sehat and Robert Jennings. 2024. "Leveraging AI for Public Health Communication: Opportunities and Risks." *Journal of Public Health Management and Practice* 30 (4): 616–18. https://doi.org/10.1097/PHH.00000000001986.
- Morgenstern, Jason D., Laura C. Rosella, Mark J. Daley, Vivek Goel, Holger J. Schünemann and Thomas Piggott. 2021. "'Al's gonna have an impact on everything in society, so it has to have an impact on public health': a fundamental qualitative descriptive study of the implications of artificial intelligence for public health." *BioMed Central Public Health* 21, 40. https://doi.org/10.1186/s12889-020-10030-x.
- Murphy, Margi. 2019. "Dr Google will see you now: Search giant wants to cash in on your medical queries." *The Telegraph*, March 10. www.telegraph.co.uk/technology/2019/03/10/ google-sifting-one-billion-health-questions-day/.
- Nicoletti, Leonardo and Dina Bass. 2023. "Humans Are Biased. Generative AI Is Even Worse." Bloomberg, June 9. www.bloomberg.com/graphics/2023-generative-ai-bias/.
- Ning, Yilin, Salinelat Teixayavong, Yuqing Shang, Julian Savulescu, Vaishaanth Nagaraj, Di Miao, Mayli Mertens et al. 2024. "Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist." *The Lancet Digital Health* 6 (11): e848–56. https://doi.org/10.1016/S2589-7500(24)00143-2.

- Office of the Privacy Commissioner of Canada. 2023. "Principles for responsible, trustworthy and privacy-protective generative AI technologies." December 7. https://priv.gc.ca/en/ privacy-topics/technology/artificial-intelligence/gd\_principles\_ai/.
- O'Neill, Rory and Mari Eccles. 2024. "Broken links and bad advice: Conversations with SARAH, WHO's virtual health aide." POLITICO, April 23. www.politico.eu/article/bad-adviceconversations-with-sarah-world-health-organization-virtual-health-aide/.
- Sallam, Malik, Nesreen A. Salim, Ala'a B. Al-Tammemi, Muna Barakat, Diaa Fayyad, Souheil Hallit, Harapan Harapan, Rabih Hallit and Azmi Mahafzah. 2023. "ChatGPT Output Regarding Compulsory Vaccination and COVID-19 Vaccine Conspiracy: A Descriptive Study at the Outset of a Paradigm Shift in Online Search for Information." *Cureus* 15 (2): e35029. https://doi.org/10.7759/cureus.35029.
- Siegrist, Michael and Alexandra Zingg. 2014. "The Role of Public Trust During Pandemics: : Implications for Crisis Communication?" *European Psychologist* 19 (1): 23–32. https://doi.org/10.1027/1016-9040/a000169.
- Smela, Beata, Mondher Toumi, Karolina Świerk, Clement Francois, Małgorzata Biernikiewicz, Emilie Clay and Laurent Boyer. 2023. "Rapid Literature Review: Definition and Methodology." *Journal of Market Access & Health Policy* 11 (1), 2241234. https://doi.org/ 10.1080/20016689.2023.2241234.
- Steerling, Emilie, Elin Siira, Per Nilsen, Petra Svedberg and Jens Nygren. 2023. "Implementing AI in healthcare the relevance of trust: a scoping review." *Frontiers in Health Services* 3, 1211150. https://doi.org/10.3389/frhs.2023.1211150.
- Swoboda, Christine M., Joseph M. Van Hulle, Ann Scheck McAlearney and Timothy R. Huerta. 2018. "Odds of talking to healthcare providers as the initial source of healthcare information: updated cross-sectional results from the Health Information National Trends Survey (HINTS)." *BioMed Central Family Practice* 19, 146. https://doi.org/10.1186/ s12875-018-0805-7.
- Thomasen, Kristen. 2023. "Regulation of AI in Canada." Symposium presented at the British Columbia Public Health Association Conference, Vancouver, BC, November 22. https://phabc.org/conference-2023/.
- WHO. 2024a. Ethics and governance of artificial intelligence for health: Guidance on large multimodal models. Geneva, Switzerland: World Health Organization. January 18. www.who.int/ publications/i/item/9789240084759.
- – . 2024b. "WHO unveils a digital health promoter harnessing generative AI for public health." News release, April 2. www.who.int/news/item/02-04-2024-who-unveils-a-digitalhealth-promoter-harnessing-generative-ai-for-public-health.