



Canadian Journal of Health Technologies

January 2024 Volume 4 Issue 1

CADTH Horizon Scan

Chatbots in Health Care: Connecting Patients to Information

Michelle Clark

Sharon Bailey



Key Messages

Why Is This an Issue?

- Artificial intelligence (AI) is increasingly being used in health care settings. Chatbots geared toward patient use are becoming more widely available, but the clinical evidence of their effectiveness remains limited.

What Is the Technology?

- AI-based chatbots are computer programs or software applications that have been designed to engage in simulated conversation with humans using humanlike language.
- Chatbots can help humans save time and allow them to focus on more high-level creative or strategic thinking by taking over more routine or repetitive tasks, such as automated customer service chats, appointments, or staff scheduling.

What Is the Potential Impact?

- Anyone with access to an internet-enabled computer or a smartphone could use these chatbots to access health information.
- Chatbots can provide patients with 24/7 access to health information, such as symptom assessment, supportive information, medication reminders, or appointment scheduling, allowing access to information when health care providers are unavailable. There appear to be trends toward efficacy and user satisfaction, but the evidence to support the clinical effectiveness of chatbots in health care is still being established.
- Existing health care chatbots are mostly free for patients to access, although some developers and health care providers charge fees to access additional features or content. Some apps may be prescribed to patients by providers. These could be covered by insurance or licensed to health care providers by the developer.

What Else Do We Need to Know?

- Ethical and data privacy issues remain top of mind when considering the widespread implementation of chatbots for patient use in health care settings.
- ChatGPT and other AI tools that were not developed specifically for health care do not necessarily provide the level of data privacy that is required of health care information. They are also trained on historical datasets and do not provide responses based on the most current clinical recommendations or health data.



Key Messages

- The development of AI-specific ethical frameworks could facilitate safer and more consistent development of AI tools in health care by preventing the misuse of AI technologies and minimizing the spread of misinformation.
- AI tools still require human oversight in terms of moderation and troubleshooting.

Purpose and Scope

The purpose of this Horizon Scan is to present health care stakeholders in Canada with an overview of information related to the use of artificial intelligence (AI) chatbots by patients in primary care, specialist care, and the provision of health and wellness information through both web-based and app-based interfaces. CADTH produced a report on [the use of AI in pre-hospital emergency health care](#) in August 2023.¹

This report is not a systematic review and does not involve critical appraisal or include a detailed summary of study findings. It is not intended to provide recommendations for or against the use of the technology and focuses only on AI chatbots in health care settings, not broader used of AI within health care.

Background: Why This? Why Now?

There is currently a shortage of health human resources. Health care providers are often stretched thin, and they may be looking for technology-based solutions, such as chatbots, to alleviate some of the administrative or other burdens placed on their time.² Patients can use chatbots when they are unable to connect with their health care provider (e.g., after business hours or when appointments are not available). Chatbots are available 24/7 without requiring human staff be available at the same time. Chatbots can answer questions more quickly than humans and handle a larger volume of requests and provide the information in a standardized format for a consistent user experience.³ This technology can also provide anonymity for people requesting information about sensitive topics.⁴

Chatbots can help patients navigate a sometimes complex health care system when used to identify available providers and to facilitate appointment scheduling. Because there are many existing and emerging AI-based health care products proposed as solutions and/or integrated into the health care system, it is important to examine their efficacy, benefits, and limitations more closely before they become embedded in the standard of care. This report provides an overview of what AI-powered chatbots for use in health care are, how AI-powered chatbots have been and are being implemented in health care settings, some evidence related to their use, and some of the issues to consider that affect the implementation of chatbots in health care.

Technology: What Is It? How Does It Work?

Chatbots are computer programs or software applications that have been designed to engage in simulated conversations with humans using natural language. Generative pretrained transformer (GPT) AI models are trained on large sets of text-based data and use that information to generate responses to questions or participate in conversations.⁵ ChatGPT is a well-known example of this type of chatbot that is optimized to produce natural, humanlike dialogue. Chatbots have been used in customer service for some time to answer customer questions about products or services before, or instead of, speaking to a human. Because these AI models are trained on existing datasets, the responses they provide will incorporate both the

factual information available to them but also any errors, inconsistencies, or biases present in the original dataset.⁵ GPT models are trained on a closed dataset and are not able to incorporate new data as it becomes available. This means, particularly in relation to medical information, that the responses they provide may be out-of-date as soon as the dataset is closed.⁵ Generalizability of AI datasets and AI algorithm bias are discussed in more detail later in this report.

There are a variety of chatbots available that are geared toward use by patients for different aspects of health. These range from chatbots that act as digital health assistants (providing services such as reminders related to health tracking, providing possible diagnoses based on inputted symptoms, and linking people with health care provider contact information) to providing cancer recovery support (such as diet and exercise information). Ten examples of currently available health care chatbots are provided in [Table 1](#).

Table 1: Examples of 10 Health Care Chatbots⁶

Chatbot name	Clinical area	Description of functionality	Accessible in Canada
Ada	General health	This chatbot has more than 13 million users and is 1 of the most popular AI-based symptom checkers available. The user's health is assessed via comparison of their symptom inputs against the large AI database that supports the chatbot. Users can share their AI assessment with their doctors. In the UK, the app can connect the user with a general practitioner.	Yes, via the Apple app store
Babylon powered by eMed	General health	Online medical consultation and health service that offers AI consultation based on a medical dataset and offers video consultation with human physicians when required.	Yes, as TELUS Health MyCare
Buoy Health	General health	The algorithm for this chatbot was trained on data from 18,000 clinical publications, 5 million patients, and 1,700 conditions. The site provides an AI-powered symptom checker and a database of articles and other information to help people determine which health issue they may be experiencing.	Yes, via website
Florence	General health	Florence chatbot works as a "personal health assistant." Users can access Florence through popular instant messaging apps. The chatbot can be used to create medication reminders, track user's health (e.g., menstrual cycle, blood pressure, weight, mood), or to help find the nearest appropriate health care professional.	Yes, via Facebook Messenger
Healthily	General health	Healthily is a version of an online symptom checker. Users can provide the chatbot with their symptoms and the bot will ask further questions and refine the responses to provide a potential diagnosis. It can also provide information directing the user to appropriate health care providers, pharmacies, or recommendations for other health apps.	Yes, via mobile app stores and web version
OneRemission	Oncology	Helps cancer patients and cancer survivors access information to help support their recovery (e.g., exercise, diet, post-cancer practices) without having to contact their doctor every time. It also provides access to online oncologists 24/7 if needed.	Unclear

Chatbot name	Clinical area	Description of functionality	Accessible in Canada
Sensely	General health	Molly, a virtual medical assistant with a human avatar, can assess the user's symptoms. Users can interact with the chatbot using speech, text, images, or video. The AI will analyze the user's symptoms and suggest a diagnosis. The colour-coded triage system is used to assign urgency to the diagnosis.	Unclear
Symptomate	General health	Users complete a 3-minute symptom assessment, via chat or voice, and are provided with a list of possible causes and recommendations for next steps.	Yes, via mobile app stores and a web version
Youper	Mental health	Monitors and aims to improve users' emotional health using CBT techniques. Users are provided with personalized meditations and the ability to monitor their emotional health and track their mood. Users communicate with a chatbot within the app that learns from the conversations and customizes the app's content to the user.	Yes
Woebot Health	Mental health	There are versions of Woebot for adults, adolescents, and maternal health. The AI-based chatbot used CBT techniques and natural language processing to converse with people experiencing anxiety or depression.	No

AI = artificial intelligence; CBT = cognitive behavioural therapy.

Place in Care: How Could This Change Care?

The most common ways AI chatbots are used in health care include:

- 24/7 availability with immediate responses which reduces or eliminates the need for patients to wait for office hours to speak to a health care providers
- automation of repetitive administrative tasks to free up time for direct patient interaction
- symptom assessment and patient triage
- standardized patient data collection that can be used to feed into an electronic health record
- quick access to critical information for users about nearby medical facilities, hours of operation, and nearby pharmacies for prescription refills
- assistance for patients in the day-to-day management of chronic conditions
- assistance determining whether a health situation is critical and requires immediate attention, and providing information about the nearest available medical facilities when required
- appointment scheduling integrated with online calendars and automatic appointment reminders
- facilitating patient requests for prescription refills, refill reminders, and submission of prescriptions to the user's pharmacy of choice
- mental health care by providing support through the chatbot or connecting the user to health care providers or peers, when appropriate.

The use of chatbots appears to be growing, particularly in the mental health space. People living with mental health conditions are able to use chatbots to “talk” about their feelings and experiences at any time they

require support and are able to address issues in the moment rather than having to wait for their health care provider to be available to speak with them.⁷ The anonymity of exchanging messages with a computer rather than a human can also increase people's comfort with disclosing sensitive or embarrassing information that they might otherwise keep to themselves.⁷ Some users might also have concerns with the use of chatbots for this purpose related to the collection and handling of personal data.

Chatbots have the potential to change access to care options for people who live in rural or remote areas and do not have easy access to health care providers in person or through telemedicine. Simple questions and information seeking can be addressed through a chatbot. People who are more comfortable with online services may choose to use a chatbot for information finding, symptom checking, or appointment booking rather than speaking with a person on the phone. Appointments for minor ailments or information gathering could potentially be directed toward an automated AI system, freeing up in-person appointments for people with more complex or urgent health issues.

People with greater digital literacy are more likely to benefit from the use of AI and find it easy to use.⁸ An access gap may be created for people who are less comfortable or familiar with technology as services shift to digital versus in-person access. People who lack reliable access to the internet or smartphones will also experience this gap in care.⁸

Regulatory Status and Availability

There is no existing specific regulatory process to authorize the use of AI-based chatbots for use in Canadian health care. In 2021, Health Canada, the US FDA, and the UK Medicines and Healthcare Products Regulatory Agency jointly released 10 guiding principles for the development of AI and machine learning in the context of medical devices.⁹ The intended use of each chatbot will determine which, if any, regulatory requirements it may need to meet to be authorized for use in health care. The use of in-house–developed AI tools or adaptations of free AI software may fall within a regulatory grey area. These principles acknowledge the growing role that AI will play in health care going forward.

As outlined in [Table 1](#), a variety of health care chatbots are currently available for patient use in Canada.

Cost and Administration

AI chatbots are becoming popular in the health care space. It is estimated that the market segment for chatbots will grow from US\$196 million in 2022 to approximately US\$1.2 billion by 2032.⁶ In terms of the prices of individual chatbots, 1 company offers the addition of a medical chatbot to a website for a fee between US\$149 to US\$400 per month.¹⁰ The cost to develop an AI-based health care chatbot from scratch may be somewhere between US\$15,000 to more than US\$100,000.¹¹ There are a variety of business models for the delivery of these apps. The freemium model provides a test or basic version of an individual service or app to users for free with some offering access to additional features for a fee.¹² Some technologies may be

prescribed by a health care provider to a patient and paid for by insurance or they may be licensed for use by health care providers from the developer.¹²

Chatbots Have the Potential to Provide Cost Savings

Most chatbots allow the user to access information about their symptoms or relevant health care–related questions 24/7, meaning they do not have to wait for health care providers to be available at a time that might be more inconvenient and/or require time off work or childcare.¹³ Chatbots can be more efficient to navigate than a traditional patient portal or telehealth system, thereby saving the user time. Chatbots can result in savings for health care providers as well by deferring some patients away from in-person appointments, which can be a cost savings to the health care system. Deferrals also free up time to see patients with more severe concerns or time to spend on other tasks.

The ability for chatbots to facilitate appointment scheduling and provide automated patient reminders can help ease the administrative burden and help to minimize the number of people who forget and do not show up for their appointments.

This allows providers to fill those empty spots with other patients or reduce clinic hours if patients will not be there.¹³ Chatbots can also contribute to savings by reducing the amount of time a live person needs to be present in a health care practice to answer phones and deal with patient queries.

What Is the Evidence?

Twelve systematic reviews and 3 scoping reviews were identified that examined the use of chatbots by patients. Further details of these reviews are available in [Table 2](#), [Appendix 2](#). The range of reported uses included chronic conditions; general use in health care; mental health; postintervention follow-up; health before, during, and after pregnancy; promotion of physical wellbeing; public health; sexual health promotion; and substance use disorder.

The chatbot interventions were determined to be effective for:

- providing information to support health behavioural changes (e.g., smoking cessation, treatment or medication adherence, and reduction in substance misuse)^{14,15}
- mental health symptom improvement (e.g., depressive symptoms, anxiety, psychosomatic disease symptoms)¹⁶⁻¹⁸
- providing users with information regarding health promotion¹⁹
- supporting increased physical activity.²⁰

The authors of 1 review reported that users found the chatbots easy to use and perceived them as a nonjudgmental way to communicate sensitive information but the authors reported mixed results when it came to feasibility, acceptability, and usability.¹⁴ Usage rates were generally high when reported,^{21,22} with high

levels of user satisfaction.^{19,23,24} Personalization and empathetic responses were reported as facilitators to chatbot use and efficacy.¹⁶ Users reported some areas for improvement even when the chatbots were well received. Some users felt the quality of information provided by the chatbot could be improved.¹⁹ Patient safety outcomes were generally not reported.²² A conclusion drawn from additional research involving chatbot users was that developers should take care to ensure they are involving the users in the design phase to create health care chatbots that are culturally appropriate for the intended audience.²⁵

The subject area is still evolving and, despite the volume of literature that exists already, authors of some of the reviews concluded that more research is required before chatbots can become part of the standard of care in many health care settings.^{3,15,18,22,26,27} There is still inconsistency and a lack of standardization in how chatbots are evaluated in health care, which makes it difficult to appropriately evaluate their effectiveness and compare outcomes between various available options.^{20,28} There was no confirmation of the usefulness of outcome measurement tools reported in most cases when they were used to evaluate a chatbot's effectiveness, and most instruments were not used in more than 1 study.²⁹

Safety

As mentioned previously, AI-based chatbots are trained using closed datasets that are not able to continuously update themselves to incorporate the most up-to-date information. This is particularly important in relation to health care, an area where clinical practice guidelines, best practices, and safety data are continuously changing. The lack of real-time updates to the content of chatbots could result in people receiving out-of-date information in response to their queries. The same can be true for human-to-human interactions; however, a health care provider does have the ability to access up-to-date information in real time, whereas an AI chatbot does not.

There have been times when chatbots have provided information that could be considered harmful to the user.

For example, Tessa, a chatbot developed to provide support to people living with eating disorders, was shut down after users discovered it was providing dieting and exercise tips to users seeking out support for the management of their eating disorders.³⁰ Tessa was not built using generative AI but was designed to deliver a specific cognitive behavioural therapy-based tool meant to prevent eating disorders. The weight loss advice that Tessa provided was not part of the data that the AI tool was meant to be trained on. The chatbot's developer has yet to address how this type of information came to be included in the chatbot's dataset.³¹ This provides a real-world example of the potential AI has to harm vulnerable groups and why human oversight remains an important factor in the incorporation of AI into health care.

User Experiences

In studies of user experiences with chatbots in health care, most user responses were positive.^{4,32-34} Adults older than 60 years who used a chatbot for health information gathering reported that the cognitive load required to use the technology was low.³³ Recommendations for improvement included stylistic

considerations³² and solving technical issues.^{32,33} Some concerns were reported regarding data privacy and security.³³

Haque and colleagues (2023)⁴ examined the user experience and characteristics of commercially available mental health apps with built-in chatbots based on reviews in US-based mobile app stores. Users expressed that personalized and humanlike interactions with the chatbots were positive. People tended to lose interest in using chatbots that provided improper or impersonal responses.⁴ The authors found that users could become overly reliant on the chatbots because they were always available. Some people seemed to withdraw from interaction with their friends and family in real life in preference for the chatbot.⁴ Chatbots lacked the ability to properly identify when users might be in crisis and require in-person care. The authors suggested that chatbot developers should give thought to when and how to limit user interaction and motivate users to use more nontechnical methods of mental health support when appropriate.⁴

Additional Considerations

Ethical Issues

Coghlan and colleagues (2023)⁷ outlined some important considerations when choosing to use chatbots in health care. Developers and professionals seeking to implement chatbots should weigh the risks and benefits by clearly defining the aim of the chatbot and the problem to be solved in their circumstances. There should be careful assessment of the problem to be solved to determine whether the use of AI or chatbots is an appropriate solution. There may be instances in which the benefits of implementation are too low or the risks are too high to justify replacing humans.⁷ The use of chatbots in health care requires an evidence-based approach. The appropriate evidence to support the safe and effective use of chatbots for the intended purpose and population should be gathered and incorporated before implementation.

Transparency in data collection and storage practices is important. In some cases, the behavioural patterns, metadata, and data entered by a user into a chatbot can be tracked and linked to their behaviours and data across the internet. Data that was thought to be anonymized could be tracked back to the specific user.⁷ It is ethically important that users understand the risks to their personal data that could result if a chatbot was hacked or the personal information provided was shared without the users' consent. Each could result in psychological harms and feelings of distrust.⁷ As such, developers have a responsibility to be transparent in their user agreements about the unique risks and harms that may be associated with their product. Developers should ensure they are following all relevant data protection laws, adequately inform the users of how their data are being used and how and where their data are being stored (e.g., locally or internationally), and determine how any data collected will be destroyed if the chatbot is shut down.⁷ Some health care providers are required to disclose to relevant authorities when they suspect that illegal acts have been reported to them (e.g., child abuse or domestic violence). Developers should determine how they would identify these kinds of disclosures and seek clarification around the legal reporting requirements that may pertain to the chatbot if a user discloses the same type of information in a forum that they believe to be confidential.⁷

To operate successfully, chatbots still require human oversight to ensure they operate the way they were designed to.⁷

There will always be glitches and technical issues that need to be fixed. Moderation may also be required but the level of moderation needs to be balanced with the associated burden on human resources.⁷ In health care, some ways AI can be used to assist providers include performing patient triage, identifying potential diagnoses based on reported symptoms, interpreting medical imaging results, and predicting the health outcomes of patients based on current symptoms and their medical history. None of the information provided by AI can be used in isolation. The analytical thinking of the health care providers is combined with the computational power of computers to incorporate the AI data into a patient's care plan in a contextually appropriate way.

Complex medical scenarios remain that cannot be adequately addressed by AI chatbots and still require human intervention to be successful, such as:³⁵

- diagnoses of rare or complex medical conditions
- surgical procedures
- counselling with human empathy
- interpretation of complex diagnostic tests
- nuanced and personalized treatment plan development.

Current iterations of generative AI systems that have not been specifically designed for health care settings do not comply with Canadian or American data protection legislation, such as the Personal Information Protection and Electronic Documents Act (PIPEDA) and the Health Insurance Portability and Accountability Act (HIPAA).^{36,37} This poses a risk of unintentional sharing of or access to personal information, including search terms and patient characteristics, by individuals outside the health care system.³⁶ The development of AI-specific ethical frameworks could facilitate safer and more consistent development of AI tools in health care by preventing the misuse of AI technologies and minimizing the spread of misinformation.^{38,39}

Before AI tools can be reliably and effectively used in health care, comprehensive plans for securely storing and managing sensitive personal information are imperative.³⁷ Furthermore, because ChatGPT is used internationally, it has faced bans in certain health care settings due to privacy concerns related to informed consent and patient data sharing.³⁷ Questions remain about the requirements of health care providers to be transparent and disclose to patients when the use of AI-generated content is being used in their care plans.³⁸

Generalizability and Algorithm Bias

Chatbots such as ChatGPT can only access the information available in the dataset they were trained on. This can introduce both statistical bias (distribution of the dataset does not accurately represent the true distribution in the population) and social bias (inequities in the data that can result in substandard outcomes for the population).⁴⁰ Middle-aged white men are the most represented group within the datasets that have so far been used to train AI chatbots.⁴¹ Health-related data for groups like Black women, Indigenous Peoples, people with disabilities, and people living in rural and remote areas are less likely to have been

proportionately included in the training datasets. This factor is considered to be a major limitation to the broad use of generalized chatbots and is a source of bias within the results the chatbots produce.⁴¹ The inclusion of communities that are underrepresented in existing health datasets in the design of AI tools can help ensure the inputs and outcomes of the algorithms accurately reflect their reality and needs. They can help to identify specific biases and gaps in the existing information that relate to themselves.⁴⁰

Generalizability and accuracy of AI could be improved by encouraging AI developers to openly share and accurately describe the data and methods used in their algorithms and the results they produce allowing for reproducibility and closer examination by researchers and other developers.⁴⁰

Algorithms and code can be shared in cases in which data cannot be shared securely or anonymously, allowing the code to be retrained on additional or more representative or setting-specific datasets.⁴⁰

Chatbots typically produce a single response based on the information contained in the dataset they are trained on, suggesting that there is a single correct answer to each question or 1 singular truth.⁴¹ Comparatively, when a user inputs the same question into a search engine, such as Google, they will be provided with pages of options that they can explore to determine which information is most relevant and useful to them depending on their personal circumstance and context. In both situations, the user should be encouraged to apply their own critical thinking skills to assess the information they have been provided.

Accessibility

Changing the way health care is delivered to rely on AI and chatbots may create some issues with who is able to access information and care. Relying more on technology means access would likely increase for some people and decrease for others. The availability and cost of smartphones and computers, as well as reliable internet access, could impact some patients' ability to access health information or health care. There may also be access considerations for people with disabilities that limit their ability to use the devices required to access the chatbots. Many chatbots rely on text-based chat, which could prove difficult to use for people with visual impairments or limitations in their ability to type. For those who cannot read or who have reading levels lower than that of the chatbot, they will also face barriers to using them.

Additionally, working knowledge of the "spoken" languages of the chatbots is required to access chatbot services. If chatbots are only available in certain languages, this could exclude those who do not have a working knowledge of those languages. Conversely, if chatbots are available in multiple languages, those people who currently have more trouble accessing health care in their first language may find they have improved access if a chatbot "speaks" their language.

Future Developments

The body of evidence will continue to grow as AI is used more often to support the provision of health care. In August 2023, a search of ClinicalTrials.gov produced 57 results of ongoing clinical trials using AI chatbots in health care. Some of the areas of study included increasing medication education and adherence, reducing vaccine hesitancy and increasing vaccine uptake, tracking and improving sleep, self-testing and reporting of communicable disease, mental health and wellbeing, diabetic foot care, and many others. The establishment of standardized usability and outcome measurement scales could aid in improving evaluation. Guidance and assessment criteria, such as the [Model for Assessment of Artificial Intelligence \(MAS-AI\)](#) in medical imaging and the [Digital Technology Assessment Criteria \(DTAC\)](#), will contribute to the continued standardization of the methods to evaluate the effectiveness of digital technologies and AI in health care.

Final Remarks

The evidence to support the effectiveness of AI chatbots to change clinical outcomes remains unclear. Chatbots cannot provide health care on their own. They require oversight from humans to ensure the information they provide is factual and appropriate. This requirement for human involvement makes it difficult to establish ability of the chatbot alone to influence patient outcomes. Researchers have recommended the development of consistent AI evaluation standards to facilitate the direct comparison of different AI health technologies with each other and with standard care. Concerns persist regarding the preservation of patient privacy and the security of data when using existing publicly accessible AI systems, such as ChatGPT. The convenience of 24/7 access to health information and the perceived confidentiality of conversing with a computer instead of a human are features that make AI chatbots appealing for patients to use.

References

1. Clark M, Severn M. CADTH health technology review: Artificial intelligence in prehospital emergency health care. *Can J Health Technol.* 2023;3(8). <https://canjhealthtechnol.ca/index.php/cjht/article/view/EN0052/EN0052>. Accessed 18 Oct 2023.
2. Ciabarra C. Healthcare workers deserve better: Can artificial intelligence help? 2023; <https://www.forbes.com/sites/forbestechcouncil/2023/09/07/healthcare-workers-deserve-better-can-artificial-intelligence-help/?sh=2698c8d0704e>. Accessed 2023 Dec 11.
3. Wilson L, Marasoiu M. The development and use of chatbots in public health: Scoping review. *JMIR Hum Factors.* 2022;9(4):e35882. [PubMed](#)
4. Haque MDR, Rubya S. An overview of chatbot-based mobile mental health apps: Insights from app description and user reviews. *JMIR Mhealth Uhealth.* 2023;11:e44838. [PubMed](#)
5. De Angelis L, Baglivo F, Arzilli G, et al. ChatGPT and the rise of large language models: The new AI-driven infodemic threat in public health. *Front Public Health.* 2023;11:1166120. [PubMed](#)
6. Meskó B. The top 10 health chatbots. 2023; <https://www.linkedin.com/pulse/top-10-health-chatbots-bertalan-mesk%25C3%25B3-md-phd/>. Accessed 2023 Oct 05.
7. Coghlan S, Leins K, Sheldrick S, Cheong M, Gooding P, D'Alfonso S. To chat or bot to chat: Ethical issues with using chatbots in mental health. *Digit Health.* 2023;9:20552076231183542. [PubMed](#)
8. Koncz A. Health equity in the AI and digital health era: Promise or peril? 2023; <https://medicalfuturist.com/health-equity-in-the-ai-and-digital-health-era-promise-or-peril/>. Accessed 2023 Oct 25.
9. Government of Canada. Good machine learning practice for medical device development: Guiding principles. 2021; <https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/good-machine-learning-practice-medical-device-development.html>. Accessed 2023 Oct 5.
10. PatientGain. Intelligent medical chatbot app for doctors. 2023; <https://www.patientgain.com/intelligent-chatbot>. Accessed 2023 Oct 5.
11. USM. How much does it cost to develop a medical chatbot? 2023; <https://usmsystems.com/how-much-does-it-cost-to-develop-a-medical-chatbot/>. Accessed 2023 Oct 5.
12. Konstantinov R. Mobile health app business models: What model to choose for revenue growth? 2023; <https://empeek.com/mhealth-business-models-how-to-choose-revenue-model-and-win/>. Accessed 2023 Nov 28.
13. Rana J. The Pros and Cons of Healthcare Chatbots. 2023; <https://www.revechat.com/blog/healthcare-chatbots/>. Accessed 2023 Oct 6.
14. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: Systematic review. *J Med Internet Res.* 2023;25:e40789. [PubMed](#)
15. Ogilvie L, Prescott J, Carson J. The use of chatbots as supportive agents for people seeking help with substance use disorder: A systematic review. *Eur Addict Res.* 2022;28(6):405-418. [PubMed](#)
16. He Y, Yang L, Qian C, et al. Conversational agent interventions for mental health problems: Systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res.* 2023;25:e43862. [PubMed](#)
17. Lim SM, Shiao CWC, Cheng LJ, Lau Y. Chatbot-delivered psychotherapy for adults with depressive and anxiety symptoms: A systematic review and meta-regression. *Behav Ther.* 2022;53(2):334-347. [PubMed](#)
18. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis. *J Med Internet Res.* 2020;22(7):e16021. [PubMed](#)
19. Balaji D, He L, Giani S, Bosse T, Wiers R, de Bruijn GJ. Effectiveness and acceptability of conversational agents for sexual health promotion: A systematic review and meta-analysis. *Sex Health.* 2022;19(5):391-405. [PubMed](#)
20. Oh YJ, Zhang J, Fang ML, Fukuoka Y. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. *Int J Behav Nutr Phys Act.* 2021;18(1):160. [PubMed](#)

21. Han R, Todd A, Wardak S, Partridge SR, Raeside R. Feasibility and acceptability of chatbots for nutrition and physical activity health promotion among adolescents: Systematic scoping review with adolescent consultation. *JMIR Hum Factors*. 2023;10:e43227. [PubMed](#)
22. Geoghegan L, Scarborough A, Wormald JCR, et al. Automated conversational agents for post-intervention follow-up: A systematic review. *BJS Open*. 2021;5(4):06.
23. Vaidyam AN, Linggonegoro D, Torous J. Changes to the psychiatric chatbot landscape: A systematic review of conversational agents in serious mental illness: Changements du paysage psychiatrique des chatbots: une revue systematique des agents conversationnels dans la maladie mentale serieuse. *Can J Psychiatry*. 2021;66(4):339-348. [PubMed](#)
24. Milne-Ives M, de Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: Systematic review. *J Med Internet Res*. 2020;22(10):e20346. [PubMed](#)
25. Chua JYX, Choolani M, Chee CYI, et al. Insights of parents and parents-to-be in using chatbots to improve their preconception, pregnancy, and postpartum health: A mixed studies review. *J Midwifery Womens Health*. 2023;68(4):480-489. [PubMed](#)
26. Bendig E, Erb B, Schulze-Thuesing L, Baumeister H. The next generation: Chatbots in clinical psychology and psychotherapy to foster mental health - a scoping review. *Verhaltenstherapie*. 2022;32(Supplement 1):64-76. <https://karger.com/ver/article/32/Suppl.%201/64/835090/The-Next-Generation-Chatbots-in-Clinical>. Accessed 2023 Aug 20.
27. Schachner T, Keller R, Wangenheim FV. Artificial intelligence-based conversational agents for chronic conditions: Systematic literature review. *J Med Internet Res*. 2020;22(9):e20701. [PubMed](#)
28. Abd-Alrazaq A, Safi Z, Alajlani M, Warren J, Househ M, Denecke K. Technical metrics used to evaluate health care chatbots: Scoping review. *J Med Internet Res*. 2020;22(6):e18301. [PubMed](#)
29. Jabir AI, Martinengo L, Lin X, Torous J, Subramaniam M, Tudor Car L. Evaluating conversational agents for mental health: Scoping review of outcomes and outcome measurement instruments. *J Med Internet Res*. 2023;25:e44548. [PubMed](#)
30. Wells K. An eating disorders chatbot offered dieting advice, raising fears about AI in health. 2023; <https://www.npr.org/sections/health-shots/2023/06/08/1180838096/an-eating-disorders-chatbot-offered-dieting-advice-raising-fears-about-ai-in-hea>. Accessed 2023 Oct 18.
31. Hoover A. An eating disorder chatbot is suspended for giving harmful advice. 2023; <https://www.wired.com/story/tessa-chatbot-suspended/>. Accessed 2023 Dec 11.
32. Dosovitsky G, Bunge E. Development of a chatbot for depression: Adolescent perceptions and recommendations. *Child Adolesc Ment Health*. 2023;28(1):124-127. [PubMed](#)
33. Wilczewski H, Soni H, Ivanova J, et al. Older adults' experience with virtual conversational agents for health data collection. *Front Digit Health*. 2023;5:1125926. [PubMed](#)
34. Shahsavar Y, Choudhury A. User intentions to use ChatGPT for self-diagnosis and health-related purposes: Cross-sectional survey study. *JMIR Hum Factors*. 2023;10:e47564. [PubMed](#)
35. Altamimi I, Altamimi A, Alhumimidi AS, Altamimi A, Temsah MH. Artificial intelligence (AI) chatbots in medicine: A supplement, not a substitute. *Cureus*. 2023;15(6):e40922. [PubMed](#)
36. Chun M. Who's liable for bad medical advice in the age of ChatGPT? 2023; <https://blog.petrieflom.law.harvard.edu/2023/06/05/whos-liable-for-bad-medical-advice-in-the-age-of-chatgpt/>. Accessed 2023 Oct 5.
37. Ashenburg N, Christian Rose C, Dayton J. Leveraging large language models (like ChatGPT) in emergency medicine: Opportunities, risks, and cautions. 2023; <https://www.acepnow.com/article/leveraging-large-language-models-like-chatgpt-in-emergency-medicine/>. Accessed 2023 Oct 5.
38. Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical considerations of using ChatGPT in health care. *J Med Internet Res*. 2023;25:e48009. [PubMed](#)
39. Guleria A, Krishan K, Sharma V, Kanchan T. ChatGPT: Forensic, legal, and ethical issues. *Med Sci Law*. 2023;258024231191829. [PubMed](#)

40. Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A. Addressing bias in big data and AI for health care: A call for open science. *Patterns (N Y)*. 2021;2(10):100347. [PubMed](#)
41. Cohen IG. What should ChatGPT mean for bioethics? *Am J Bioeth*. 2023;23(10):8-16. [PubMed](#)

Appendix 1: Methods

Note that this appendix has not been copy-edited.

Literature Search Strategy

An information specialist conducted a literature search on key resources including MEDLINE, Embase, the Cochrane Database of Systematic Reviews, the International HTA Database, the websites of Canadian and major international health technology agencies, as well as a focused internet search. The search approach was customized to retrieve a limited set of results, balancing comprehensiveness with relevancy. The search strategy comprised both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. Search concepts were developed based on the elements of the research questions and selection criteria. The main search concept was AI chatbots. The search was completed on August 14, 2023, and limited to English-language documents published since January 1, 2020. Regular alerts updated the database literature searches until October 2, 2023.

Study Selection

One author screened the literature search results and reviewed the full text of all potentially relevant studies. Studies were considered for inclusion if the intervention was chatbots or AI conversational agents used in health care settings. Conference abstracts and grey literature were included when they provided additional information to that available in the published studies.

Peer Review

A draft version of this bulletin was reviewed by 1 reviewer with expertise in AI and informatics.

Appendix 2: Systematic and Scoping Review Details

Note that this appendix has not been copy-edited.

Table 2: Systematic and Scoping Reviews of AI Chatbots in Health Care

Author and year	Topic area	Type of review, number of included studies, outcomes	Main findings
Aggarwal et al. (2023) ¹⁴	Health behavioural changes	<ul style="list-style-type: none"> • SR • 15 studies • Outcomes <ul style="list-style-type: none"> ◦ Health behavioural outcomes ◦ Feasibility ◦ Usability ◦ Acceptability ◦ Engagement 	<ul style="list-style-type: none"> • Results of the included studies showed high efficacy for promoting healthy lifestyles, smoking cessation, treatment or medication adherence, and reduction in substance misuse • The included studies reported mixed results relating to feasibility, acceptability, and usability • Users reported that chatbots provided them with a nonjudgmental space to communicate sensitive information
Chua et al. (2023) ²⁵	Preconception, pregnancy, and postpartum health	<ul style="list-style-type: none"> • SR • 15 studies • Outcomes <ul style="list-style-type: none"> ◦ Clinical effectiveness ◦ User experience • Acceptability and satisfaction 	<ul style="list-style-type: none"> • Chatbots can help improve health by providing information and psychological support • Developers should engage with users to create culturally appropriate chatbots that meet the needs of the specific communities they are targeting
Han et al. (2023) ²¹	Nutrition and physical activity health promotion	<ul style="list-style-type: none"> • Systematic scoping review with adolescent consultation • 5 studies • Outcomes: <ul style="list-style-type: none"> ◦ Clinical effectiveness ◦ Feasibility ◦ Acceptability 	<ul style="list-style-type: none"> • The usage rate of the chatbots was above 50% in 3 of the 5 included studies • 1 study reported positive change in health-related outcomes • During consultation, adolescents identified issues with the design of the chatbots that have not been described in the existing literature such as putting these apps through a formal review of the accuracy of their information and provide them with a 'stamp of approval'

Author and year	Topic area	Type of review, number of included studies, outcomes	Main findings
He et al. (2023) ¹⁶	Mental health	<ul style="list-style-type: none"> • SR and MA of RCTs • 32 studies • Outcomes: <ul style="list-style-type: none"> ◦ Clinical effectiveness ◦ Quality of life ◦ Positive and negative affect 	<ul style="list-style-type: none"> • Statistically significant positive short-term effects were demonstrated for depressive symptoms, generalized anxiety symptoms, specific anxiety symptoms, mental disorder symptoms, psychosomatic disease symptoms, and negative affect • Long-term effects were not statistically significant • Personalization and empathetic response were reported as facilitators of efficacy
Balaji et al. (2022) ¹⁹	Sexual health promotion	<ul style="list-style-type: none"> • SR and MA • 31 articles • Outcomes: <ul style="list-style-type: none"> ◦ Clinical effectiveness ◦ Acceptability 	<ul style="list-style-type: none"> • The authors determined the included studies provided preliminary support for the effectiveness of chatbots for promoting sexual health, including treatment adherence • High levels of user satisfaction were reported • Users felt the quality of the information provided could be improved
Bendig et al. (2022) ²⁶	Mental health	<ul style="list-style-type: none"> • Scoping review • 6 studies • Outcomes NR 	<ul style="list-style-type: none"> • The authors concluded that chatbot technology is still experimental in the area of psychology and psychotherapy • Identified studies were most often pilot studies without randomized comparison groups
Lim et al. (2022) ¹⁷	Mental health <ul style="list-style-type: none"> • Depressive and anxiety symptoms 	<ul style="list-style-type: none"> • SR and meta-regression • 11 RCTs • Outcomes: <ul style="list-style-type: none"> ◦ Depressive symptoms 	<ul style="list-style-type: none"> • MA resulted in statistically significant improvement in depressive symptom scores with a medium effect size
Ogilvie et al. (2022) ¹⁵	Substance use disorder	<ul style="list-style-type: none"> • SR • 6 studies • Outcomes <ul style="list-style-type: none"> ◦ Efficacy ◦ Usability 	<ul style="list-style-type: none"> • The quantity of research in the field is limited and more study is needed on the usefulness of chatbots for SUD • None of the studies considered whether chatbots were better than other existing digital interventions

Author and year	Topic area	Type of review, number of included studies, outcomes	Main findings
			<ul style="list-style-type: none"> A reduction in substance use was observed in 2 studies
Wilson et al. (2022) ³	Public health	<ul style="list-style-type: none"> Scoping review 32 studies Outcomes: <ul style="list-style-type: none"> Clinical effectiveness Adherence 	<ul style="list-style-type: none"> The evidence for efficacy for prevention and intervention is limited Mental health and COVID-19 were the public health domains identified
Geoghegan et al. (2021) ²²	Postintervention follow-up for physical health interventions <ul style="list-style-type: none"> Elective orthopedic surgery Orthopedic trauma surgery Surgical intervention for varicose veins Breast cancer Uteroscopy Medical management or hypertension, asthma, or cancer 	<ul style="list-style-type: none"> SR 10 studies Outcomes: <ul style="list-style-type: none"> Clinical outcomes Adherence Patient experience Patient safety 	<ul style="list-style-type: none"> All chatbots in the included studies were used on mobile phones Engagement rates ranged from 31% to 97% None of the included studies examined patient safety outcomes Further study is required before chatbots are integrated into routine clinical care
Oh et al. (2021) ²⁰	Promotion of physical activity, healthy diet, and weight loss	<ul style="list-style-type: none"> SR 9 studies Outcomes: <ul style="list-style-type: none"> Changes in physical activity, sedentary behaviour, diet, body weight Feasibility Acceptability Safety User satisfaction 	<ul style="list-style-type: none"> 5 of the 7 studies of physical activity found chatbots to be promising to help people increase their movement The authors could not make definitive conclusions about the effectiveness of chatbots for physical activity, diet, or weight management
Vaidyam et al. (2021) ²³	Serious mental illness <ul style="list-style-type: none"> Anxiety disorder Bipolar disorder Major depressive disorder Schizophrenia spectrum disorders 	<ul style="list-style-type: none"> SR 7 studies Outcomes: <ul style="list-style-type: none"> Clinical effectiveness Feasibility User engagement Acceptability Adherence 	<ul style="list-style-type: none"> The authors found there was a lack of standardized measures to evaluate effectiveness and facilitate comparison Users were generally satisfied with their experience Pediatric patients and people with schizophrenia and bipolar disorder were not included in the identified studies
Abd-Alrazaq et al. (2020) ¹⁸	Mental health	<ul style="list-style-type: none"> SR and MA 12 studies Outcomes: <ul style="list-style-type: none"> Depression 	<ul style="list-style-type: none"> Chatbots have the potential to improve mental health Insufficient evidence was identified to conclude their

Author and year	Topic area	Type of review, number of included studies, outcomes	Main findings
		<ul style="list-style-type: none"> ○ Anxiety ○ Positive and negative effect ○ Subjective psychological wellbeing ○ Psychological distress ○ Stress ○ Acrophobia ○ Safety 	<ul style="list-style-type: none"> • effect is clinically important • Further studies are required
Milne-Ives et al. (2020)²⁴	General use in health care	<ul style="list-style-type: none"> • SR • 31 studies • Outcomes: <ul style="list-style-type: none"> ○ Clinical effectiveness ○ Usability ○ Satisfaction ○ Improvement in health care provision ○ Resource implications 	<ul style="list-style-type: none"> • Overall, evidence was mostly positive or mixed • Usability and satisfaction performed well • 23 of 30 studies reported positive or mixed effectiveness • Qualitative interviews in the included studies included feedback about the limitations of the chat programs
Schachner et al. (2020)²⁷	Chronic conditions <ul style="list-style-type: none"> • Alzheimer disease • Asthma • COPD • Depression/anxiety • Diabetes • Glaucoma • Heart failure • JIA • Parkinson/dementia • Psoriasis 	<ul style="list-style-type: none"> • SR • 10 studies • Outcomes: <ul style="list-style-type: none"> ○ Clinical effectiveness ○ User engagement ○ Patient satisfaction ○ Algorithm performance 	<ul style="list-style-type: none"> • The authors determined the literature to be scarce and of poor quality • Further research is required to evaluate and compare the effectiveness and quality of chatbots developed for chronic conditions

AI = artificial intelligence; COPD = chronic obstructive pulmonary disease; JIA = juvenile idiopathic arthritis; MA = meta-analysis; NR = not reported; RCT = randomized controlled trial; SR = systematic review; SUD = substance use disorder.

Authors: Michelle Clark, Sharon Bailey

Contributors: David Rotenberg, Chief Analytics Officer, Centre for Addiction and Mental Health

ISSN: 2563-6596

Disclaimer: The information in this document is intended to help Canadian health care decision-makers, health care professionals, health systems leaders, and policy-makers make well-informed decisions and thereby improve the quality of health care services. While patients and others may access this document, the document is made available for informational purposes only and no representations or warranties are made with respect to its fitness for any particular purpose. The information in this document should not be used as a substitute for professional medical advice or as a substitute for the application of clinical judgment in respect of the care of a particular patient or other professional judgment in any decision-making process. The Canadian Agency for Drugs and Technologies in Health (CADTH) does not endorse any information, drugs, therapies, treatments, products, processes, or services.

While care has been taken to ensure that the information prepared by CADTH in this document is accurate, complete, and up-to-date as at the applicable date the material was first published by CADTH, CADTH does not make any guarantees to that effect. CADTH does not guarantee and is not responsible for the quality, currency, propriety, accuracy, or reasonableness of any statements, information, or conclusions contained in any third-party materials used in preparing this document. The views and opinions of third parties published in this document do not necessarily state or reflect those of CADTH.

CADTH is not responsible for any errors, omissions, injury, loss, or damage arising from or relating to the use (or misuse) of any information, statements, or conclusions contained in or implied by the contents of this document or any of the source materials.

This document may contain links to third-party websites. CADTH does not have control over the content of such sites. Use of third-party sites is governed by the third-party website owners' own terms and conditions set out for such sites. CADTH does not make any guarantee with respect to any information contained on such third-party sites and CADTH is not responsible for any injury, loss, or damage suffered as a result of using such third-party sites. CADTH has no responsibility for the collection, use, and disclosure of personal information by third-party sites.

Subject to the aforementioned limitations, the views expressed herein are those of CADTH and do not necessarily represent the views of Canada's federal, provincial, or territorial governments or any third-party supplier of information.

This document is prepared and intended for use in the context of the Canadian health care system. The use of this document outside of Canada is done so at the user's own risk.

This disclaimer and any questions or matters of any nature arising from or relating to the content or use (or misuse) of this document will be governed by and interpreted in accordance with the laws of the Province of Ontario and the laws of Canada applicable therein, and all proceedings shall be subject to the exclusive jurisdiction of the courts of the Province of Ontario, Canada.

The copyright and other intellectual property rights in this document are owned by CADTH and its licensors. These rights are protected by the Canadian *Copyright Act* and other national and international laws and agreements. Users are permitted to make copies of this document for non-commercial purposes only, provided it is not modified when reproduced and appropriate credit is given to CADTH and its licensors.

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to Requests@CADTH.ca



Canada's Drug and
Health Technology Agency