The Development and Evaluation of a Frequently-Asked Question (FAQ) Answering Chatbot for an Online Mental Wellness Program for People with Chronic Disease

by

Chikku Preci Sadasivan

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Medicine

University of Alberta

ABSTRACT

Introduction: Chatbots are computer programs that interact with humans through natural language conversations. Chatbots are on the rise due to numerous benefits like convenience, 24/7, personalized support, and they show promise in healthcare to support people living with chronic diseases. While chatbots, like other digital health interventions, are promising tools for chronic disease management, they also face challenges such as limited user adoption and low engagement. One way to address these challenges is to involve patients in development. Objectives: The purpose of these studies was to report the non-technical (e.g. non-software development related) approaches for chatbot creation while examining patient engagement in these approaches and to develop the FAQ chatbot and evaluate its acceptability, usability, and user engagement through a multi-method approach.

Methods: In Chapter 3, our team conducted a scoping review following the framework proposed by Arksey and O'Malley. Nine electronic databases were searched in July 2022. Studies were selected based on our inclusion/exclusion criteria. Data were then extracted, and patient involvement was assessed. In Chapter 4, a FAQ-answering chatbot ("Liv") was developed from May 2022 to February 2023 and evaluated through a multi-phased, multi-method design. Liv was then deployed on an online mental wellness program and evaluated for acceptability, usability, and user engagement.

Results: In Chapter 3, 16 studies were included in the review. We report several approaches to chatbot development, assess patient involvement where possible, and reveal the limited detail available on reporting patient involvement in the chatbot implementation process. In Chapter 4, Liv was deployed for 120 days, and there were 259 conversations with Liv, with 175 instances of active engagement (back-and-forth user-chatbot interaction). Engagement was highest during the first 30 days of deployment. The chatbot had a resolution rate of 33%. Liv's usability was evaluated with an overall score of 50.8 (below average), with successes in specific areas, including its navigation and ease of use. Findings from qualitative interviews included comments on Liv's personality, the convenience of knowing the chatbot would escalate, and experiences with Liv and her improvement over time.

Conclusion: FAQ chatbots may be an engaging way to provide patient support in online mental wellness programming. Including patients in development may improve chatbot acceptability, usability, and user engagement. Future work that prioritizes patient engagement in the development and builds upon conversation log data to create a more advanced bot is warranted.

PREFACE

This thesis is an original work by Chikku Preci Sadasivan. The research project of which this thesis is a part received research ethics approval from the University of Alberta Research Ethics Board. Project name "The Equitable Mental Wellness Programming for Older Adults with multiple chronic conditions (EMPOwer) Randomized Controlled Trial", Pro00122568, November, 22, 2022.

Chapter 3 of this thesis has been published as C. Sadasivan, C. Cruz, N. Dolgoy, A. Hyde, S. Campbell, M. McNeely, E. Stroulia, and P. Tandon, "Examing Patient Engagement in Chatbot Development Approaches for Health Lifestyle and Mental Wellness Interventions: Scoping Review," Journal of Participatory Medicine, vol. 15, e45772. I was responsible for conducting the scoping review and composing the manuscript. C. Cruz and S. Campbell assisted with conducting the scoping review. N. Dolgoy, A. Hyde, M. McNeely, and E. Stroulia were involved with manuscript edits. P. Tandon was the senior author involved in concept formation and manuscript edits.

Chapter 4 of this thesis is in the process of being published as C. Sadasivan, A. Hyde, E. Johnson, E. Stroulia, and P. Tandon, "The development and evaluation of a frequently-asked question answering chatbot for an online mental wellness program." Co-authors listed have contributed to either: i) conception or design of work [CS, ES, PT], ii) acquisition, analysis, or interpretation of data for the work [CS, AH, EJ, PT], or iii) drafting the work or revising it critically for important intellectual content [CS, AH, ES, PT].

DEDICATION

This thesis is dedicated to anyone living with a chronic condition. Thank you for volunteering your time to take part in the development and evaluation of our chatbot and contributing to this thesis, and for being open to trying new things. You have all helped me grow so much over my master's, and I look forward to continued growth and learning as I enter the medical community.

ACKNOWLEDGEMENTS

I will begin by thanking the members of my supervisory committee. To my primary supervisor, Dr. Puneeta Tandon: thank you for taking me on as a master's student. The opportunities you have provided me over the past two years have made me grow into who I am today, and you have impacted my personal and professional self. You have challenged and put me in multiple positions to learn new skills. Your encouragement, guidance, and kindness have been key to my growth and success over the last two years. To Dr. Ashley Hyde: thank you for the academic support and personal guidance you have provided me. You have passed on so much knowledge and have given me many practical tips to improve my academic writing. To Dr. Eleni Stroulia: thank you for your wisdom, enthusiasm, and time supporting me with this work. Before conducting this thesis work, your supervision through my reading course was essential to my learning.

Thank you to the Alberta SPOR Support Unit for introducing me to patient-oriented research and funding 12 months of my master's work. Moreover, thank you to the patients who were involved in my master's research for your curiosity, encouragement, and commitment to my research.

Thank you to the graduate students in the Tandon Lab: Christofer Cruz, Ana Teresa Limon, Ashley Hyde, Emily Johnson, Makayla Watt, Shaina Corrick, and Serena Isley. I am lucky to now call you my friends, in addition to being colleagues. You have made my master's experience successful and entertaining. I have enjoyed making memories together, learning from you, and I am forever grateful for your contributions to my research work and development.

Thank you to my family, my mom and dad, Kavitha Chandrasenan and Chandran Sadasivan, my brother Chandu Sadasivan, my girlfriend Aashna Mann, and the newest addition - my puppy Babu Sadasivan-Mann. Your unconditional love, knowledge, and words of motivation have helped me throughout my education and my master's.

Finally, I would like to thank the University of Alberta's MD Program for granting me a one-year deferral to finish my master's studies before beginning medical school. This deferral allowed me to learn and grow as a person and professional, which will be crucial to succeeding as a medical student and physician.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Brief Introduction

Chatbots are computer programs that use artificial intelligence to interact with humans through natural language conversations in text or speech [1]. They were first introduced in the 1960s with the development of Eliza, a simple chatbot that simulated psychotherapist conversation [2, 3]. Over time, chatbots have evolved and can now perform various tasks, including customer service, education, personal assistance, marketing, sales, entertainment, companionship, and information retrieval [1-3]. Chatbots have become a prominent feature of many digital applications, notably intelligent voice assistants like Amazon's Alexa, Apple's Siri, Microsoft's Cortona, and Google Assistant [4].

As with chatbots, digital applications or technologies have rapidly increased in recent years. This is particularly evident in healthcare delivery, with the emergence of digital health applications, which include mobile health (mHealth) interventions, wearable devices, personalized medicine, and artificial intelligence to improve healthcare services and health outcomes [5]. Increasingly, chatbots are integrated into these digital health applications to provide timely, effective, and scalable user support solutions [6, 7]. For example, chatbots have been used for patient education [8-11], providing medication management [12-14], checking symptoms [15-18], promoting healthy lifestyle changes [19-23], and delivering mental health support [24-28].

Digital health interventions show promise in supporting the management of chronic diseases [29, 30] like heart disease, cancer, diabetes, arthritis, and lung disease [31]. Often caused by behaviours like tobacco use, poor nutrition, limited physical activity, and excessive alcohol consumption, chronic diseases impact rates of death and disability, with significant adverse effects on patient quality of life [31, 32]. In Canada, one in three individuals lives with at least one major chronic disease [33]. Chronic diseases also contribute to seven of the top ten leading causes of death worldwide, with an estimated 41 million deaths annually [34]. This number continues to grow yearly.

Lifestyle interventions, programs designed to promote healthy lifestyle habit formation, can help prevent the onset and progression of chronic diseases by targeting behaviour change [35]. Specifically, they can positively impact an individual's physical and mental health by promoting healthy lifestyle choices through physical activity, dietary modifications, smoking cessation, and stress management [36-40]. Unfortunately, digitally delivered lifestyle interventions face limited user adoption, low engagement, and high attrition rates that limit intervention effectiveness [41-44].

Researchers have explored ways to address these challenges in digital health applications, and one proposed solution is chatbots [6, 45, 46]. When well-trained, chatbots can act as virtual assistants that can provide immediate, efficient, and scalable 24/7 user support, potentially improving the user experience [6, 7, 47-49]. Researchers also acknowledge chatbots' limitations, such as low user uptake, user intent classification errors, and gaps between user expectations and chatbot abilities [50, 51]. One approach to mitigate the gap between user expectations and chatbot functionality is involving patients or end-users in the development process, focusing on user-centred design [52-54]. User-centred design prioritizes user (patient) needs and perspectives in designing digital applications [55]. Involving patients in the development process can help developers identify critical requirements that could be overlooked, which is vital in maximizing application usability, engagement, and intervention effectiveness [53, 56]. Further research is necessary to understand what is known about the involvement of patients in the chatbot development process, and to explore the acceptability and usability of chatbot technology to support the digital delivery of lifestyle and mental wellness interventions in individuals living with chronic disease.

1.2 Problem Statement and Purpose of the Thesis

The thesis seeks to enhance the delivery of an online mental wellness program for people with chronic disease [57] by:

(i) Reviewing what is known in the literature about the non-technical (e.g. non-software development related) steps for developing chatbots, focusing on patient-oriented approaches.

(ii) Developing the knowledge bank of an FAQ-based chatbot informed by patient needs and perspectives.

(iii) Assessing the acceptability, usability, and user engagement with the FAQ-based chatbot embedded within a 12-week online mental wellness program (EMPOWER).

Previous versions of the 12-week online mental wellness programming demonstrated promising results for improving mental wellness and quality of life [57]. The current version of the program (EMPOWER) is based on previous research. In those studies, a gap has been the need for patients to reach out via email to a research team member. With the potential of chatbots to improve user support, we sought to evaluate that further as part of this thesis. Looking at the literature, limited studies have investigated the usefulness of FAQ-based virtual assistant chatbots to support the delivery of digital health interventions [58]. To our knowledge, health chatbot research has focused on chatbots that directly administer an intervention or chatbots that streamline healthcare service provision [19, 21, 24, 59]. Additionally, there has been a gap in that limited research has outlined the chatbot development process, specifically with patient involvement in design and implementation.

The research reported in this thesis aims to address these research gaps by developing and evaluating a chatbot with a framework-guided approach informed by patient needs and perspectives. The thesis includes a scoping review to describe the approaches to optimize chatbot development and to examine the level of patient engagement in these approaches. It also includes a multi-method development study that documents the chatbot development process and evaluates its acceptability, usability, and user engagement. This thesis aims to collaborate with patients in developing and testing a chatbot as a supportive tool for a 12-week online mental wellness program.

1.3 Objectives

Scoping Review (Chapter 3): The overall aim of the scoping review presented in Chapter 3 was to examine the development of chatbots that engage in two-way natural language interaction through voice- or text-based input to aid the delivery of interventions that support healthy eating, physical activity, and mental wellness. The specific objectives of this study were: (1) to report the non-technical (e.g. non-software development related) approaches for chatbot creation and (2) to examine the level of patient engagement in these reported approaches.

Development and Evaluation Study (Chapter 4): This development and evaluation study presented in Chapter 4 aimed to work with patients to develop a functional frequently asked question (FAQ) answering chatbot to support patients as they navigate through a 12-week mental wellness program. The specific objective of this study was to develop the FAQ chatbot and to evaluate its acceptability, usability, and user engagement through a multi-method approach. The multi-method approach includes both qualitative and quantitative components:

Qualitative Component: To understand the chatbot's acceptability, usability, and user engagement by examining patient perspectives, perceptions, and experiences through patient interviewing and chatbot conversation log analysis.

Quantitative Component: To examine the chatbot's usability through the chatbot usability questionnaire and to measure user engagement with the chatbot by analyzing conversation log data.

1.4 Hypothesis

Scoping Review Hypothesis: We hypothesize that there will be limited literature describing the non-technical approaches of chatbot development or the level of patient involvement in these approaches.

Chatbot Evaluation Hypothesis: We hypothesize that patients involved in the 12-week EMPOWER study will actively engage with the chatbot and view the chatbot positively in terms of its acceptability and usability.

CHAPTER 2: LITERATURE REVIEW

2.1.1 What are chatbots?

Chatbots are artificial intelligence programs designed to simulate conversation with human users in natural language through text or speech [1]. Also known as virtual assistants, conversational agents, text bots, interactive agents, talkbots, messaging (IM) bots, and conversational artificial intelligence (AI), they all have the same goal: to enable natural language communication between a human user and a machine [60]. Chatbots can be integrated into existing websites and applications through chat windows or deployed on messaging applications like Facebook Messenger and WhatsApp.

Chatbots have various applications, including customer service, sales, personal assistance, healthcare, education, banking and finance, human resources, entertainment, and travel [61]. The popularity of chatbot development and usage resides in the benefits they offer users and providers. Chatbots have the potential to provide users with convenient, 24/7, personalized, and engaging support. For providers, they have the potential benefit of improving customer service, saving costs on personnel, increasing scalability, and improving efficiency [6, 7, 47-49]. With advances in artificial intelligence and natural language process technology, chatbot usage is anticipated to continue to increase [62].

2.1.2 Brief History of Chatbots

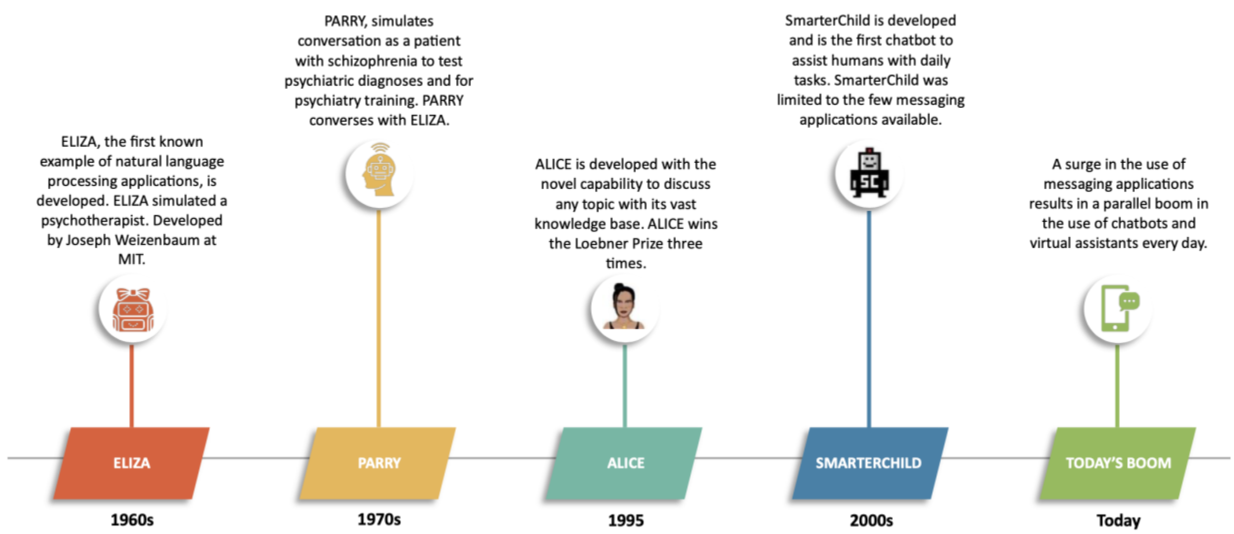
To better understand the potential of chatbot technology in healthcare, it is important to review the history of chatbot development **(Figure 2.1).** The first chatbot, ELIZA, was created by Joseph Weizenbaum at the Massachusetts Institute of Technology in the 1960s [2, 63].

ELIZA used a template-based response mechanism to mimic the conversational approach of a non-directive psychotherapist [60, 63]. ELIZA served as a model for chatbots that followed because of its success in simulating human conversation [2]. Another early chatbot, PARRY, simulated conversation with a patient with schizophrenia, and was used to test psychiatric diagnoses and as a tool to train young psychiatrists to communicate with patients with schizophrenia [1, 2]. Interestingly, these early chatbots had applications in mental health and represent the first recorded developments in chatbot history.

Though they didn't have an application to healthcare, the chatbots that followed PARRY and ELIZA represented notable technological advancements. In 1995, ALICE (Artificial Linguistic Internet Computer Entity) was the first online chatbot that used pattern matching to discuss any topic on the web [1]. ALICE won the Loebner Prize thrice due to its novelty [1, 60]. This prize is an annual competition in artificial intelligence research for the best human-like computer program [1, 60]. Another significant milestone in chatbot history was the development of SmarterChild in the early 2000s [64]. SmarterChild was the first chatbot able to retrieve information from databases to help humans with daily tasks such as providing weather information, news, and sports scores [1]. However, SmarterChild's availability was limited by the few messaging applications available at the time, like American Online (AOL) and Microsoft Messenger (MSN) [64].

In recent years, there has been a tremendous increase in the use of mobile devices, social media and messaging applications, with many interacting with chatbots daily [60]. Since the introduction of ELIZA, more than 50 years ago, texting and messaging have become a routine means of communication worldwide, with over 6 billion people using SMS-capable mobile phones [65]. The growth in mobile devices, the internet, social networking and messaging applications created the optimal environment for chatbot development and usage [2, 60]. Today, for most people, communicating through short, typed-out interactions is comfortable, natural, and convenient [60]. Many people use chatbots daily while interacting with popular intelligent voice assistants such as Amazon's Alexa, Apple's Siri, Microsoft's Cortona, and Google Assistant [4]. Chatbots are becoming increasingly sophisticated and are anticipated to evolve with increasing development, research, and implementation [60].

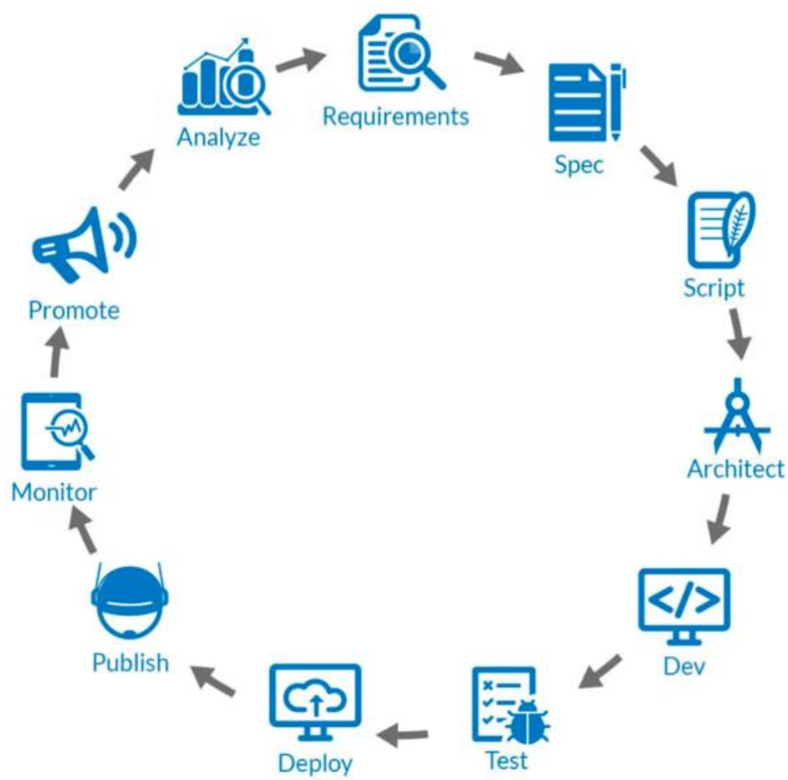
Figure 2.1. A timeline of the brief history of chatbots.



2.1.3 Contemporary Approaches to Chatbot Development

As chatbot usage is on the rise, many companies and individuals are looking to develop chatbots for their applications. Today, chatbots can be developed rapidly through chatbot-building platforms like Drift, Intercom, IBM Watson Assistant, FreshChat, and Zendesk, which do not require coding [66]. For more complex applications, developers can build chatbots from scratch via coding and using existing natural language processing libraries. Still, implementation can take longer and presents a challenge for beginners or non-technical users [67, 68]. In Pérez-Soler et al.'s study on choosing chatbot development tools, they suggested that when a chatbot lives in a narrow domain (like serving on a specific, closed website), it may be simpler to train the chatbot with training phrases or intents specific to this narrow domain [69]. Regardless of the platform or approach used for chatbot implementation, experts suggest that consulting a bot development framework for guidance makes the process easier and faster while maximizing functionality [67]. According to Cameron et al., the current best practice framework in chatbot development is the Chatbot Development Life Cycle **(Figure 2.2),** which recognizes the cyclic and iterative nature of chatbot development [67, 70].

Figure 2.2. Chatbot Development Life Cycle [70].



The Chatbot Development Life Cycle is a framework that encapsulates the chatbot building process [70]. This cycle was derived from Beerud Sheth and his chatbot software company's experience building hundreds of advanced chatbots to assist large and small businesses [70]. It consists of 11 repeated steps in a cyclic development process [70]. The 11 steps included in the cycle are described below.

1. Requirements: End-user needs and issues the chatbot can address are identified [70].
2. Spec: Key features and functionalities are identified, and the chatbot is given a short and long description for development [70].
3. Script: Conversational scripts to represent user interactions are built [70].
4. Architect: The chatbot's front-end interface and back-end design is created [70].
5. Develop: Bot developers iterate between coding, building, and testing the conversational interface [70]. Note: this step does not require coding when using a bot platform to build.
6. Test: The bot's conversation interface and code are tested on the emulator and messaging platform [70]. This is intertwined with step 5.
7. Deploy: The bot is deployed in its host environment. While in the host environment, the bot will need additional monitoring and developer support [70].
8. Publish: The bot can be submitted to a messaging platform or app store for approval. Messaging platforms have specific guidelines and criteria to meet [70].
9. Monitor: After the bot is deployed or published, the user conversation is monitored and conversational script is improved [70].
10. Promote: The bot is advertised and marketed to introduce the bot to new users. For industrial use, chatbots can be advertised on cross-platform bot stores. This step functions to bring the bot greater traffic [70].

11. Analyze: Bot performance, conversation logs, and usage metrics are tracked and reviewed [70].

2.1.4 Chatbot Applications in Healthcare: Potential Opportunities

Along with the rapid increase in the use of chatbots in various industries, chatbot use in healthcare has also increased. Recent systematic review evidence supports chatbots' potential to reduce the burden on primary care physicians and assist patients with caring for their health and wellness [71]. Some current examples of how chatbots are used in healthcare include patient education [8-11], providing medication management [12-14], checking symptoms [15-18], promoting healthy lifestyle changes [20-23, 72], and delivering mental health support [24-28].

A number of reviews have explored using chatbots to promote physical activity, healthy diet, weight loss, and mental health [24, 45, 59, 72-74]. A systematic review of chatbots in these areas identified that most health chatbots focus on mental health, followed by physical activity, with few focusing on dietary modification or weight loss [72]. As a result, Oh et al. could only review 9 articles that met their inclusion criteria, with 4 being randomized controlled trials and 5 being quasi-experimental studies [72]. Their review observed that chatbots could significantly improve physical activity behaviors, such as moderate to vigorous exercise increases, and diet-related behaviours, such as reductions in red and processed meat consumption [72]. In mental health, chatbots have been used to deliver therapy (ex. CBT), train providers and patients (ex. social skill development for People with Autism Spectrum Disorder), and screen patients for various mental illnesses [24]. In Abd-Alrazaq et al.'s scoping review of 53 studies, it was found that most mental health chatbots focus on anxiety and depression, with fewer targeting schizophrenia, dementia, phobic disorders, stress, eating disorders, obsessive-compulsive disorder, and bipolar disorder [24].

2.1.5 Chatbot Applications in Healthcare: Potential Challenges

While primary evidence from reviews in this area reveals the potential that chatbots hold to improve healthcare; chatbots also face numerous challenges. First, systematic reviews have been unable to synthesize findings across studies and make definitive conclusions on chatbot efficacy due to limited consistency in reporting of outcomes [24, 59, 72, 73]. This is due to limitations in the tools used for reporting chatbot-related outcomes and measuring participant engagement, acceptability, and usability with chatbot technology [19]. Additionally, the lack of randomized controlled trials compounds the inconsistent reporting of chatbot-related outcomes and makes interpreting results across different studies, their systems and designs even more challenging [24, 75]. Another limitation, as discussed in more detail in Chapter 3, is that there is limited published detail describing the chatbot development process, including the patient involvement in this process [76]. Many chatbot systems have been developed without regard for theoretical foundations to guide the design approach or the chatbot's conversational strategies for chatbots that target behavior change [77]. The incomplete reporting of formative stages and the lack of theoretical foundations in chatbot development brings difficulty for future researchers to replicate and implement successful chatbots. Lastly, there are additional concerns about the application of chatbots in healthcare, which include privacy issues, sharing of sensitive information, liability, and reliability [78]. The data stored on chatbot vendor's servers are a security concern and a potential area where breaches may occur [79]. Failure to provide reliable and accurate information to patients can reduce their trust in novel technology and make them anxious to use it [79, 80].

2.2.1 What are Chronic Diseases?

Chatbots show potential in the management of chronic diseases. The World Health Organization (WHO) defines chronic diseases as long-term (lasting >1 year), non-contagious, progressive illnesses that require medical attention and impact activities of daily living [34]. Risk factors for chronic disease include background (age, sex, education, genetics), harmful behaviors (tobacco use, alcohol consumption, physical inactivity, unhealthy diet), social/economic conditions (poverty, unemployment, family composition), environmental factors, and culture [81, 82]. Often resulting from the interaction of genetic, physiological, environmental and lifestyle factors, the most prevalent chronic diseases include cardiovascular diseases (ex., heart attacks, strokes), cancers, chronic respiratory diseases (ex., chronic obstructive pulmonary disease, asthma), and arthritis [34, 83].

2.2.2 Prevalence & Impact of Chronic Disease

Chronic diseases increasingly contribute to global mortality rates, rising from 28 million annually in 1990 to 41 million deaths in 2022 [34]. In Canada, a third of the population lives with at least one chronic disease, with approximately 44 % living with two or more chronic conditions [84]. The rise of chronic disease in Canada can be attributed to several factors, including an aging population, increased life expectancies, and the influence of risk factors [85]. The prevalence and burden of chronic diseases are expected to continue to increase [86, 87].

Chronic disease has a widespread impact, impacting an individual patient's health and well-being [78], and our healthcare system [88]. In addition to being a leading cause of mortality, chronic disease can negatively affect health-related quality of life (HRQoL) [89]. HRQoL is "the value assigned to the duration of life as modified by the impairments, functional states, perceptions and social opportunities influenced by disease, injury, treatment or policy." [90]. People with chronic disease often experience a diminished HRQoL, marred by chronic pain, fatigue, and other symptoms that can impact their personal and professional lives [91, 92]. Further, they often experience emotional distress caused by uncertainty, loss of control and the overwhelming nature of multiple healthcare appointments [91].

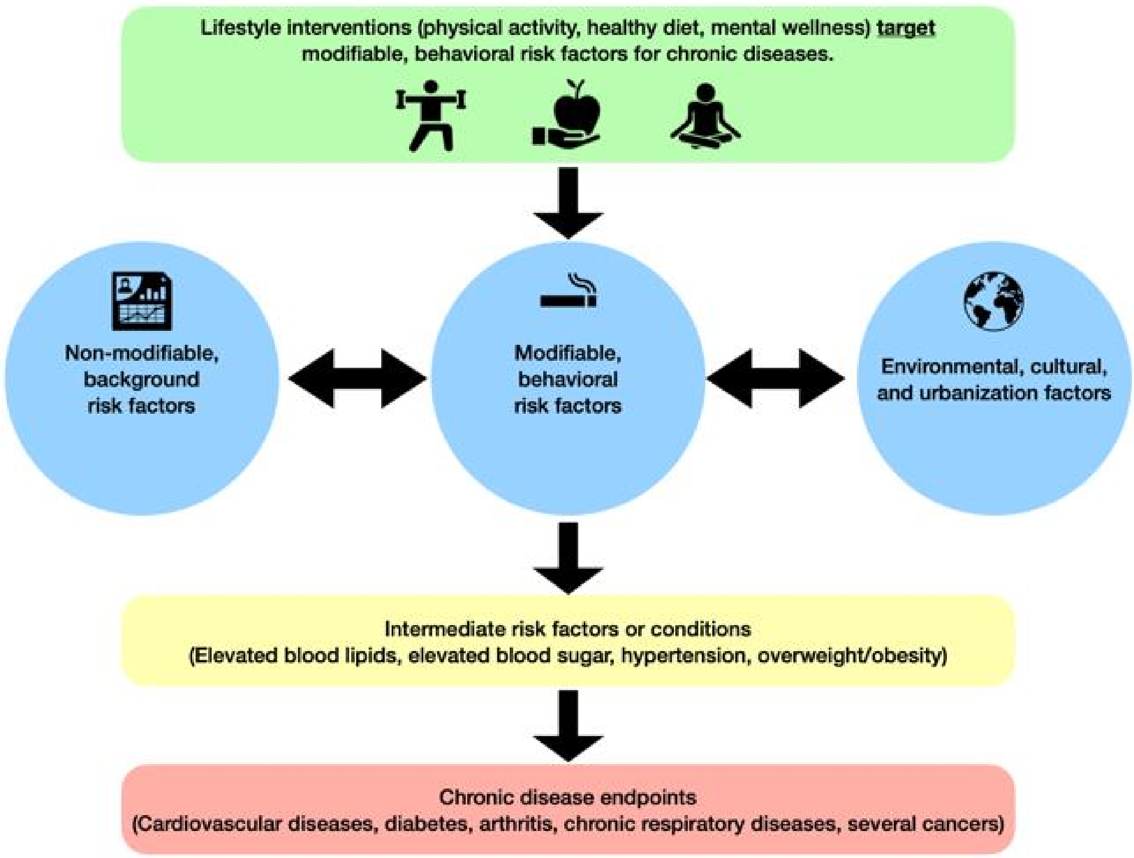
Chronic diseases also significantly impact the healthcare system and society at large [88]. They account for over $190 billion per year in both direct medical costs ($68 billion) and indirect costs ($122 billion) related to lost income and national productivity [88, 93]. Moreover, healthcare providers can find it challenging and emotionally straining to manage multiple chronic conditions with limited resources [94].

2.2.3. Approaches to Chronic Disease Management

Traditionally, chronic diseases are managed through pharmacologic approaches to control symptoms and slow disease progression [95, 96]. However, given that modifiable, behavioral risk factors contribute to many chronic diseases[81], lifestyle interventions show promise in modifying harmful behaviors through health education, diet modification, psychotherapy, behavioural counselling, and physical active components , and ultimately can result in better patient outcomes **(Figure 2.3)** [97, 98]. These interventions can encourage users to engage in goal setting, confront barriers to healthy living, stress management, cognitive restructuring, relapse prevention, and provide individual support [99]. They can be delivered via various modalities, including face-to-face, telephone, web-based programming, and digital modules [97]. Individuals participating in comprehensive lifestyle modification programs experience rapid, clinically significant, and significant improvements in health, quality of life, and psychosocial outcomes [39, 100-106]. Some examples of lifestyle changes that result from intervention include ample consumption of fruits and vegetables, limiting sodium and sugar intake, limiting excessive caloric intake, maintaining regular physical activity through walking and cycling, limiting television watching, smoking cessation, and moderating alcohol consumption [107].

Although lifestyle interventions to improve health outcomes are critical to comprehensive chronic disease management, these interventions are limited by a lack of patient adherence [108-110]. Several barriers that limit patient adherence include lack of motivation, social pressures, time constraints, physical limitations, lack of knowledge or understanding around the importance of the intervention, negative thoughts, socioeconomic constraints, and health literacy gaps [108]. With improved adherence, treatment effectiveness, health outcomes, and chronic disease burden are improved [108].

Figure 2.3. Interaction between chronic disease risk factors and lifestyle interventions [82].



2.3.1 Digital Health, Opportunities, and Challenges

Today, many lifestyle interventions are delivered through digital technologies, including artificial intelligence (AI), mobile health applications, wearable devices, telemedicine, and health information technology [111]. Digital health solutions offer numerous advantages, including increased healthcare accessibility and convenience for patients in remote and underserved areas [112], improved provider efficiency [113], enhanced patient engagement in self-management [114, 115], and healthcare system cost savings [116].

While digital health technologies show promise, there are several notable challenges. For example, digital health can perpetuate health inequity due to limited access to technology and information security concerns [111, 117-119]. Additionally, digitally delivered interventions face adherence challenges like limited user adoption, low engagement, and high attrition rates that can limit effectiveness [41-44]. Experts suggest that user-centred design principles may address the adherence issue by including patients in the digital health development process [53, 54, 120].

2.3.2 User-centered design

"User-centered design" (UCD) is a term from Computer Science that refers to designing products, systems, or services with input from end-users, allowing them to influence the design process [55]. UCD aims to prioritize end-user needs by understanding their goals, behaviours, and characteristics [55]. Incorporating end-users into the design process can result in more intuitive, engaging, and effective products with increased user satisfaction, improved usability, and higher adoption rates [56].

2.3.3 Patient-oriented research: definitions and importance

UCD and patient-oriented research (POR) share similar goals, focusing on prioritizing end-users' or patients' needs and preferences in the design process [53, 55, 56, 121]. In healthcare, many UCD approaches (ex., interviews, surveys, observational studies, usability testing, iterative prototyping) have been identified as key parts of the POR approach [55]. POR places patients (individuals with personal experience of a health issue and their informal caregivers, including family and friends) at the center of the research process, actively engaged in design, implementation, and dissemination of research [121, 122].

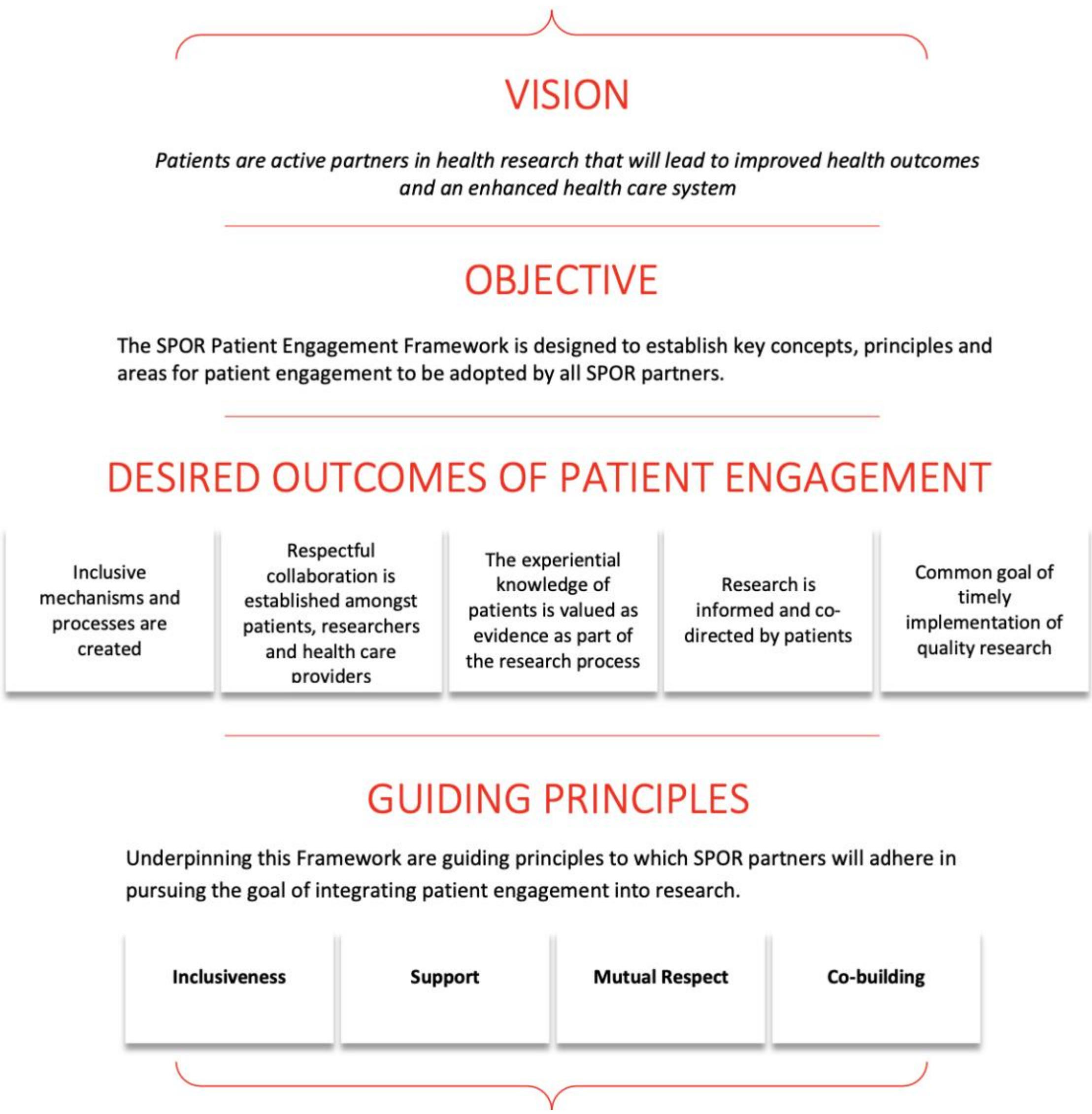
Developers and researchers often overlook the inclusion of patients in developing digital health applications [53, 123]. Several opportunities exist throughout the development process to include end-users or patients [53, 124]. For example, patients can participate in focus groups for prototype conceptualization, co-design to enhance the intervention and participate in iterative testing [124]. Though it can be challenging, including patients in the design and testing phases of digital health interventions has benefits in increasing the usability, uptake, and effectiveness of digital health interventions [53, 56, 125, 126].

2.3.4 Guidance for conducting patient-oriented research

The significance of POR is undeniable, though its practical application may need to be clarified. Several frameworks have been established to assist researchers in this process, like the Strategy for Patient-Oriented Research (SPOR) framework **(Figure 2.4)** developed by the Canadian Institutes of Health Research (CIHR) [127]. The SPOR framework has contributed to advancing patient-oriented research practices in Canada [127, 128], with positive influence reported across several research studies [128]. The SPOR framework provides guidance on why patient contribution is important and what it looks like in practice [127]. It outlines the variety of roles that patients have in research, including as committee members, researchers, contributors to design, and supporters of participant-friendly studies [127]. The framework includes four guiding principles (inclusiveness, support, mutual respect, and co-build) to help researchers pursue the goal of integrating patient engagement in research [127].

Additionally, the framework recognizes the importance of evaluating patient engagement, highlighting conditions for success in POR [127]. One tool to assess engagement is the Guidance for Reporting Involvement of Patients and Public (GRIPP2) short-form checklist [129]. The GRIPP2 checklist is a short, five-item checklist designed based on SPOR principles to enhance the quality of patient and public involvement (PPI) reporting in health research [129]. Finally, the framework outlines considerations for compensating patient partners for their time, expertise, and contributions [127].

Figure 2.4. SPOR patient engagement framework dashboard [127].



2.4 Summary and Rationale

In summary, chronic diseases are a growing problem with impacts on individuals and society at large [88]. Digital health interventions show promise in addressing the lifestyle factors contributing to chronic diseases [97, 98, 114, 115]. Chatbots are on the rise and are becoming a part of digital health interventions because they provide immediate, efficient, and scalable 24/7 user support, improving the user experience [6, 7, 47-49]. Chatbots show promise in healthcare delivery, lifestyle management, and mental wellness support [24, 45, 59, 72-74]. While chatbots have numerous benefits, like other digital health technologies, they face challenges such as limited user adoption and low engagement [41-44]. To build better patient-friendly chatbots, researchers suggest using a patient-oriented approach to development [53, 54, 120]. Researchers and developers for health chatbots have yet to explore the process of patient-oriented chatbot development in detail. Additional work on the importance of including the patient voice in chatbot development is warranted to build accessible and effective technology for patients.

CHAPTER 3: EXAMINING PATIENT ENGAGEMENT IN CHATBOT DEVELOPMENT APPROACHES FOR HEALTHY LIFESTYLE AND MENTAL WELLNESS INTERVENTIONS: A SCOPING REVIEW

Authors: Chikku Sadasivan, Christofer Cruz, Naomi Dolgoy, Ashley Hyde, Sandra Campbell, Margaret McNeely, Eleni Stroulia, Puneeta Tandon

3.1. ABSTRACT

Background

Chatbots are growing in popularity as they offer a range of potential benefits to end-users and service providers. Our scoping review aimed to explore studies that use two-way chatbots to support healthy eating, physical activity, and/or mental wellness interventions.

Objective

Our objectives were to report the non-technical (e.g. non-software development related) approaches for chatbot development and to examine the level of patient engagement in these reported approaches.

Methods

Our team conducted a scoping review following the framework proposed by Arksey and O'Malley. Nine electronic databases were searched in July 2022. Studies were selected based on our inclusion/exclusion criteria. Data was then extracted and patient involvement was assessed.

Results

16 studies were included in this review. We report several approaches to chatbot development, assess patient involvement where possible, and reveal the limited detail available on reporting of patient involvement in the chatbot implementation process.

Conclusions

The approaches reported in this review and the identified limitations hold promise to guide the inclusion of patient engagement and the improved documentation of engagement in the chatbot development process for future healthcare research. Given the importance of end-user involvement in chatbot development, it is our hope that future research will more systematically report on chatbot development, and more consistently and actively engage patients in the co-development process.

Keywords: chatbots, virtual assistants, patient involvement, patient engagement, co-development

3.2. Introduction

Growing evidence supports the use of digital technology in healthy eating, physical activity, and mental wellness interventions. Several systematic reviews on these digital health interventions (DHIs) have identified their promise in managing chronic disease [29, 130-134]. Specifically, DHIs have proven impacts on reducing risk factors for chronic disease [132, 133] by increasing physical activity, reducing body mass index [134] and improving patient psychosocial well-being [131]. Further, DHIs can help overcome barriers to access to mental health support for individuals with chronic conditions [130]. While these DHIs are useful in vulnerable chronic disease populations [29, 30], they face challenges, including limited user adoption, low engagement, and high attrition rates [41-44].

Chatbots are artificial intelligence programs that converse with humans through natural language in text or speech [1]. There is a growing body of evidence that the integration of chatbots into DHIs may provide support [19, 53, 135-137] by increasing patient engagement [135], intervention adherence[135], and the acceptability and efficacy of lifestyle/wellness interventions [53, 136-138]. Additionally, chatbots offer a range of potential benefits to end-users and service providers, most notably allowing for more scalable, cost-efficient, and interactive solutions [1].

Although developments in artificial intelligence (AI) and computer science have improved the ability of chatbots to mimic human agents, the acquisition of a relevant dataset with which to train chatbots remains challenging. User-centered design with public and patient involvement (PPI) may offer a potential solution [139-141]. By engaging key stakeholders, PPI can help produce better quality interventions relevant to end-user needs [139], resulting in benefits such as increasing intervention acceptability, effectiveness and sustainability [140]. Drawing on evidence across other digital healthcare innovations, the proposed benefits of PPI fundamentally include the development of interventions that are both usable by and relevant to patients [140]. Recognizing the limited data available to guide the role of PPI in digital health innovation, experts have called for the meaningful involvement of patients from the beginning of the development process to allow for the co-creation of relevant, valuable, and acceptable digital health solutions [141].

This scoping review aimed to map the literature on studies using chatbots to engage in two-way natural language interaction (voice- or text-based input) to aid the delivery of healthy eating, physical activity, and/or mental wellness interventions. The specific objectives of this review were: (1) to report the non-technical (e.g. non-software development related) approaches for chatbot creation and (2) to examine the level of patient engagement in these reported approaches. Although the technical software development steps are essential to creating chatbots, this review focused on the non-technical approaches for chatbot development as these are less explored and more likely to involve patient participation. To our knowledge, this is the first scoping review to systematically explore these objectives.

3.3. Methods

3.3.1. Study Design

This scoping review was conducted using the framework proposed by Arksey and O'Malley [142], and later refined by Levac and colleagues [143]. The Arksey and O'Malley framework consists of the following five steps: (1) identify a research question, (2) identify relevant studies, (3) select studies, (4) chart the data, and (5) summarize and report the results [142]. Two research questions guided the review:

1. Outside of the technical software development processes, what approaches are described for the development of chatbots that support healthy eating, physical activity, and/or mental wellness interventions?

2. What is the extent of patient engagement in these approaches?

3.3.2. Study Team

Our multidisciplinary study team included two graduate student researchers (CS, CC), a health sciences librarian (SC), two post-doctoral fellows with backgrounds in clinical care and scoping reviews (ND, AH), a professor of Medicine (PT), a professor of Physiotherapy (MM) and a professor of Computing Science (ES).

3.3.3. Search Strategy

A health sciences research librarian (SC) was consulted to develop a search strategy that used concepts from our research questions. The search strategy **(Figure 3.1)** included a combination of subject headings and keywords, including health, chatbots, and lifestyle/wellness components. Searches were adjusted appropriately for each database. Nine electronic databases were searched in July 2022 including OVID Medline, OVID EMBASE, OVID PsycINFO, EBSCO CINAHL, Scopus, IEEE Explore, Proquest Dissertations and Theses Full Text, Cochrane Library and Prospero. No publication date limit was applied on the search, as the literature of chatbots and virtual conversation agents is naturally self-limiting. After conducting the search, the results were imported into *Covidence* systematic review management software and duplicates were removed [144]. *Covidence* is a "web-based collaboration software platform that streamlines the production of systematic and other literature reviews" [144].

Figure 3.1. Search strategy used for OVID PsycINFO database.

#	Searches
1	(chatbot* or "im bot" or "im bots" or "instant message bot*" or "conversational agent*" or "virtual agent*").mp.
2	*"Diets"/
3	*"Health Promotion"/
4	*"Intervention"/
5	*"Physical Activity"/
6	"Nutrition"/
7	"Weight Loss"/
8	"Sedentary Behavior"/
9	(lifestyle* or health* or medic* or nursing or nurse* or disabilit* or elder* or "senior citizen*" or patient* or exercise or "physical activit*" or motivational).mp.
10	2 or 3 or 4 or 5 or 6 or 7 or 8 or 9
11	1 and 10

3.3.4. Eligibility Criteria

Included publications were those written in English and published in peer-reviewed journals. Included studies all had an intervention supporting healthy eating, physical activity, and/or mental well-being. All studies required a chatbot which communicated with users through a two-way natural language interaction. Inclusion criteria for participants consisted of adolescents (age

> 10 years old) as defined by the World Health Organization [145] or adult populations. Studies were excluded if they involved additional technologies or chatbot service delivery beyond the scope of this review (i.e. embodied conversation agents, humanoid/social robots, wearable technology, internet of things (IoT), virtual avatars, interactive voice assistants, or chatbots delivering therapy to clients). Studies were also excluded if they only described an intervention but did not conduct or test one. Chatbots designed to replace a therapist's role were excluded, as were papers that did not present original results (i.e. reviews and protocol papers). RCTs were included in recognition that they often contain valuable insights into the development process, particularly when the authors did not publish a formative manuscript.

3.3.5. Study Selection

Titles and abstracts of the retrieved articles were reviewed independently by two researchers (CS, CC) based on the inclusion and exclusion criteria described above. Both reviewers met throughout the title and abstract screening stage to discuss and resolve conflict through consensus. A third reviewer (ND, AH) was consulted for consensus. The remaining articles advanced to the full-text screening stage. Excluded articles were tagged with reasons for exclusion derived from our exclusion criteria. After independent full-text screening, both reviewers met to resolve any inclusion/exclusion and exclusion reason conflicts. Inter-rater reliability was assessed using Cohen’s Kappa [146]. For the included articles, an additional literature search was carried out using the involved authors, chatbot details, and reference lists to determine whether previous formative papers that described the chatbot development had been published.

3.3.6. Data Extraction

One reviewer (CS) extracted the data from included articles using a standardized Microsoft Excel form. General and specific data were extracted, including: author, publication year, journal, study setting, study design, sample size, participant demographics (age, sex, and chronic disease where applicable), intervention type, chatbot type, chatbot development approaches, and assessment of patient involvement in development.

Patient involvement was assessed using the Guidance for Reporting Involvement of Patients and Public (GRIPP2) short-form checklist [129]. The GRIPP2 checklist was applicable for our objectives as it was designed to enhance the quality of patient and public involvement (PPI) reporting in health technology assessment and health research [129], and because it could be used retrospectively to measure quality of PPI reporting in publications and reports [147]. **Figure 3.2** depicts the GRIPP2 checklist as we used it to assess PPI in chatbot development. The GRIPP2 awards points across five items that describe public engagement and involvement.

Figure 3.2. How the GRIPP2 reporting checklist was used to grade patient and public involvement in chatbot non-technical development. Adapted from reference by Staniszewska et al [147].

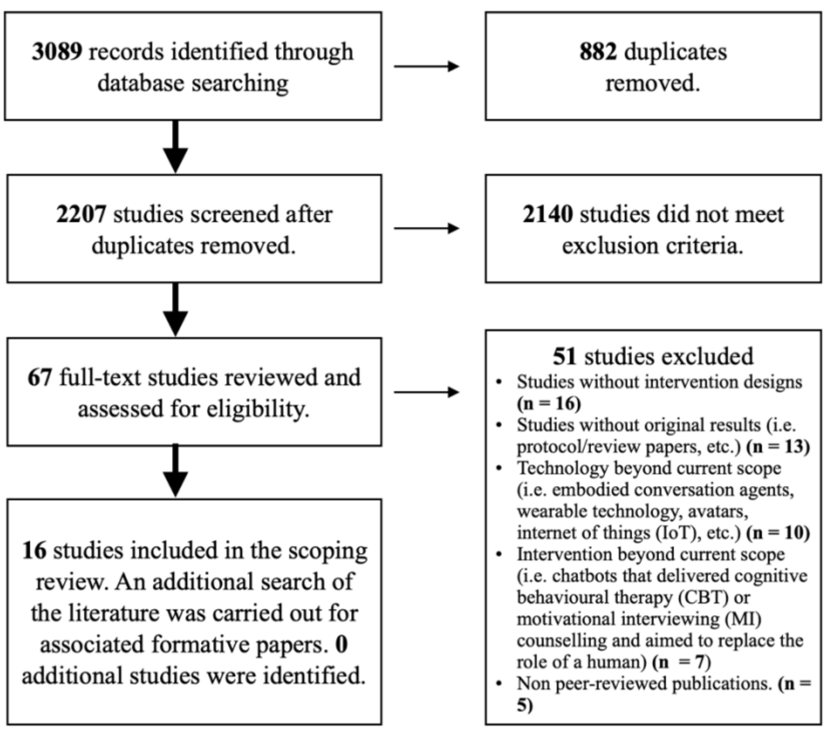
Section and topic	Specifics for engagement in chatbot-related development
1: Aim	Report the aim of PPI in chatbot development
2: Methods	Provide a clear description of the methods used for PPI in chatbot development
3: Study results	Outcomes - Report the results of PPI in chatbot development, including both positive and negative outcomes
4: Discussion and conclusions	Outcomes - Comment on the extent to which PPI influenced chatbot development overall. Describe positive and negative effects
5: Reflections/critical perspective	Comment critically on chatbot development, reflecting on the things that went well and those that did not, so others can learn from this experience

3.4. Results

3.4.1 Search Results

Figure 3.3 shows the search results; 3089 publications were retrieved from the database searches, and 882 duplicates were removed, leaving 2207 studies to screen. At the title and abstract screening stage, there was "fair" agreement between reviewers (Cohen's Kappa = 0.309, Proportionate Agreement = 0.967). After completing the title and abstract screening, 2140 publications were removed as they did not meet the inclusion criteria. Reading the full text of the remaining 67 publications resulted in a further 51 publications being excluded, with the exclusion reasons documented in **Figure 3.3**. At the full-text review stage, there was "almost perfect" agreement (Cohen's Kappa = 0.843, Proportionate Agreement = 0.941). In total, 16 publications were included in this review.

Figure 3.3. Prisma flow diagram of included and excluded studies.



3.4.2. Description of Included Studies

Table 3.1 shows the description of the included studies and their chatbot interventions. The included studies were conducted in four countries, with 50% (8/16) of the studies conducted in Canada [148-155]. Six studies were conducted in Switzerland [156-161]. One study was conducted in Saudi Arabia [162] and one study was conducted in Korea [163]. The majority of the studies (14/16) were conducted in a healthcare setting [148-160, 163], with the remaining two studies in a computing science setting [161, 162]. All but one of the included studies [151] were published in 2020 or later.

Table 3.1. Descriptive summary of included studies, chatbots, and their development.

Study and Country	**Study type**	**Chatbot intervention**	**Approaches for development**	**Identified development approaches**	**Patient engagement (GRIPP2)**
Alghamdi et al. (2021) Saudi Arabia	Randomised controlled trial	Text-based nutrition chatbot for patients with celiac disease.	Literature review of existing health behavior change models. Investigated pros and cons of each model to guide development of a health behavior change model to structure the chatbot's content. Interviews with expert users (from patient population diagnosed with celiac disease 4+ years ago, patient's parent, dietitian supervising patient for 4+ years, gastroenterologist treating celiac disease patient for 4+ years) Questionnaires for patients with celiac disease to understand symptoms and technology use preferences.	Literature review Patient interviews Collaboration with knowledge experts	Unable to assess.
Davis et al. (2020) Switzerland	Non-randomised experimental study	Text-based exercise and nutrition chatbot.	Development outsourced to a software company; did not report any steps taken for development.	None identified.	Unable to assess.
Dhinagaran et al. (2021)	Feasibility study	Text-based exercise, nutrition, and wellness chatbot for patients with diabetes.	Needs assessment conducted in an earlier publication. Literature review of systematic reviews and clinical guidelines for evidence-based content development to develop content.	Literature review Patient interviews	Unable to assess.

			After a 4 week pilot feasibility study, conducted follow-up interviews to understand patient views of the chatbot and to gain ideas for improvement.		
Figueroa et al. (2021) Switzerland	User design study	Text-based exercise chatbot.	Qualitative interviews during prototype testing to assess opinions and knowledge of chatbots as personal health coaches, technology use, digital literacy, and privacy considerations of chatbots in general. Wizard of Oz procedure. Participants completed a 20 minute text-messaging conversation with a simulated chatbot. Chatbot prototype testing. Participants texted the prototype for 10–20 minutes. Directly after the testing period, participants had a semi-structured interview via videoconference regarding the chatbot's ease of use, usefulness, humanness, and sustainability. Co-design workshop for participants to take part in development of ideas for chatbot use and design. These workshops were held over Zoom and ideas were visualized on Google Jamboard.	Patient interviews Wizard of oz procedure Prototype testing Co-design workshops	Met criteria on GRIPP2 checklist points 2, 4, and 5. Provided a clear description of the methods used for PPI, commented on how PPI influenced the study, and on successful and unsuccessful aspects of the study relating to PPI.
Gabrielli et al. (2021) Canada	Proof-of-concept study, mixed-methods	Text-based wellness chatbot.	Intervention design. The intervention, targets, and components were defined to specify clinically relevant effects on users and to refine the intervention components. This was done by a team of three clinical psychologists, two users, and behavior change experts. Preliminary testing. A proof-of-concept implementation of the digital intervention and chatbot to examine engagement and effectiveness with a convenience sample of university students.	Collaboration with knowledge experts Prototype testing	Unable to assess.
Gabrielli et al. (2020) Canada	Pilot, co-design study	Text-based wellness chatbot.	Co-design workshop. The students used and commented on a prototyped session of the chatbot intervention to collect their needs and preferences on the following: the chatbot's look and feel, the type of content and duration of the session, their unmet expectations regarding the prototype, and suggested improvements. Feasibility test. This formative study aimed to assess the perceived value of the coaching intervention and to check the user experience with intervention to refine content.	Co-design workshops	Met criteria on GRIPP2 checklist point 2. Provided a clear description of the methods used for PPI.

Greer et al. (2019) Canada	Randomised controlled trial	Text-based wellness chatbot for patients with cancer.	Literature review of the Stress and Coping theory and the Broaden-and-Build theory of positive emotion and focused on the teaching and practice of eight positive psychological skills. Created lessons based on this review for the chatbot to deliver. Interviews and focus groups as formative work to refine content for the chatbot format and inform adaptation for delivery to a young user-base with a shared experience of cancer treatment.	Literature review Patient interviews	Unable to assess.
Issom et al. (2021) Switzerland	Usability study	Text-based exercise, nutrition, and wellness chatbot for patients with sickle cell disease (SCD).	Literature review of evidence-based knowledge of SCD self-management, in addition to consulting World Health Organization' handbooks on how to implement text-based mHealth interventions to help with dialogue design.	Literature review	Unable to assess.
Krishnakumar et al. (2021) Canada	Non-randomised experimental study	Text-based exercise and nutrition chatbot for patients with Type 2 diabetes mellitus.	Literature review to develop a lesson plan of the program. This was based on the American Association of Diabetes Educator's AADE7 self-care behaviors.	Literature review	Unable to assess.
Larbi et al. (2022) Switzerland	Usability study	Text-based exercise chatbot.	Literature review of behavior change interventions. Summarized and briefly reported four steps in development: strategy planning, design, implementation, and testing. As part of strategy planning, Psychology and Public Health experts were interviewed. Also stated that the development of the prototype involved three steps: requirement analysis, concept development, and implementation. Reporting did not go into any further detail.	Literature review	Unable to assess.
Maenhout et al. (2022) Switzerland	Development pilot study	Text-based exercise, nutrition, and wellness chatbot.	Intervention planning through a scoping review of literature, conducting focus groups, and consulting online chat threads for a youth helpline. Focus groups addressed: content preferences, design preferences, questions that the chatbot would be asked, and answers that were expected from the chatbot. Intervention optimization through conducting a log data analysis during pretesting. A prototype of the chatbot was developed and pretested by the target users. The prototype was developed based upon guidance from phase 1 focus	Literature review Patient interviews Prototype testing	Met criteria on GRIPP2 checklist point 2.

			groups. Conversation logs were closely monitored to refine and fine-tune the chatbot. A question list was formed at the end of this prototype testing phase, 37 new (and practical) questions originated that were not covered in the chat threads and focus groups.		
Maher et al. (2020) Canada	Proof-of-concept study	Text-based exercise and nutrition chatbot.	Did not report how the chatbot was developed; the methods section described how the pilot study was conducted.	None identified.	Unable to assess.
Pecune et al. (2021) Switzerland	Non-randomised experimental study	Text-based nutrition chatbot.	Literature review of persuasive systems, recommender systems, and food related experiments. Collected a food database by regrouping the 40 ingredients that people most frequently cook and eat for dinner. This data was collected from hundreds of participants through questionnaires. Completed a pilot study to determine what the critical elements are for recipe recommendation systems. Also, completed this quasi-experimental study to understand the efficacy of different chatbot characteristics with the target end user group.	Literature review	Unable to assess.
Piao et al. (2020) Korea	Usability study	Text-based exercise chatbot.	Needs assessment through online surveys to assess daily routines of office workers (the target group). This was used to determine daily activities that were measurable and easy to execute. These became a part of the goal setting in the intervention. Chatbot design was guided through review of literature and to determine a theoretical model for the chatbot's basis: the habit formation model. Conducted this formative usability test prior to the RCT below to identify issues and make revisions.	Literature review Prototype testing	Unable to assess.
Piao et al. (2020) Canada	Randomised controlled trial	Text-based exercise chatbot.	Literature review of extrinsic and intrinsic reward systems. Steps for development were documented in the usability study described above.	Literature review	Unable to assess.
To et al. (2021)	Non-randomised experimental study	Text-based exercise chatbot.	Development was outsourced for technical development by SmartAI. Did not report if the research team was involved in any other steps for development.	None identified.	Unable to assess.

Canada					

3.4.3. Study Design and Interventions

Three of the included studies were randomized controlled trials (RCTs) [151, 155, 162], four were non-randomised experimental studies [152, 154, 156, 161], three were user-design and development studies [150, 157, 160], three were usability studies [158, 159, 163], one was a feasibility study [148], and two were proof-of-concept studies [149, 153].

15 of the 16 included studies reported the sample size; sample sizes ranged from 18 to 116 participants [154, 157]. Participant age ranged from 12 – 69 years, with most participants being less than 50 years old. When a specific chronic disease group was described, populations included patients with celiac disease [162], diabetes [148, 152], cancer [151], and sickle cell disease [158]. Where reported, the inclusion of female participants ranged from 31.4% to 100% [157]. Five studies involved an exercise intervention [154, 155, 157, 159, 163]. Three studies included a mental wellness intervention for healthy coping, life skill coaching, and positive psychology skill building [149-151]. Two studies evaluated a nutrition intervention [161, 162]. The remaining interventions combined exercise, nutrition, and/or mental wellness components [148, 152, 153, 156, 158, 160]. Across all reviewed articles, the chatbots communicated with users through text.

3.4.4. Study Findings

There were several approaches used to guide the development and training of chatbots. In three of the included studies, the non-software development approaches for chatbot development were not documented; therefore, no approaches were identified [153, 154, 156]. Thirteen studies reported approaches taken for chatbot development, with most studies reporting multiple approaches [148-152, 155, 157-163]. In four of the 13 studies, patients were engaged as knowledge experts or participants in co-design workshops [149, 150, 157, 162]. In six of the 13 studies, patients were involved in the study as research participants and, as part of the study outcomes, were invited to share their views through interviews, prototype testing, and the Wizard of Oz (WoZ) procedure [148, 151, 157, 160, 162, 163]. Ten of these 13 studies used a literature review, an approach that did not involve patients [148, 151, 152, 155, 158-163]. Notably, seven of the 16 included studies were already at a more advanced stage of chatbot development, focusing on evaluating interventions and usage instead of focusing on the development process itself [151, 152, 154-156, 161, 162]. Within these studies, researchers often briefly described their overall approaches but did not go into detailed steps or explain why those steps were considered important. This did range from study to study. In one non-randomised experimental study, it was reported that development was outsourced to a software company without further details regarding the process [156]. In contrast, one RCT effectively described the formative work their team did working with patients to refine content through interviews and focus groups [151]. However, the degree of utilization and success of the development strategy was not discussed [151]. Although we searched the literature for formative papers that preceded the included papers, no additional studies were identified using this approach (Figure 3). These non-technical development approaches are listed and described in more detail below.

3.4.5. Collaboration with Patient and/or Clinician Partners as Knowledge Experts

During the early stages of chatbot planning, two studies consulted experts for chatbot development [149, 162]. In both studies, patient partners were recognized as knowledge experts and included as part of the research team [149, 162]. In the study with a nutrition chatbot for a celiac disease patient group, patients were recognized as experts alongside healthcare professionals, including dietitians and gastroenterologists [162]. In the mental wellness study, a team of three clinical psychologists took part in chatbot intervention development and content refinement alongside two users and a group of behavior change experts; this iterative process was used to adapt the chatbot's intervention program, and audiovisual content to user needs through a clinical lens [149].

3.4.6. Co-Design Workshops

Two studies used co-design workshops to allow patients to creatively engage in the development of content ideas, chatbot design, chatbot style elements, and chatbot use [150, 157]. One study invited participants to collaborate and develop ideas together with the research team over Zoom (an online communication platform) by visualizing ideas on Google Jamboard software (a virtual whiteboard for idea sharing) [157]. Another study invited patients to use a prototyped session with the chatbot to collect their needs, content preferences, stylistic ideas, and suggestions for improvements [150].

3.4.7. Interviews with Patients

In five studies, patient interviews were conducted beforehand to guide chatbot development by exploring patient needs, perceptions, and experiences with chatbot use and healthy living [148, 151, 157, 160, 162]. In one study, interviews were administered during prototype testing and analyzed qualitatively [157]. Another study conducted this formative work through focus groups and interviews to collect information from young adults treated for cancer, the target end-user population [151]. This information was then used to guide chatbot content development within a patient-centered lens. Follow-up interviews were conducted after interventions or chatbot exposure [148, 160]. Questionnaires and surveys were also used in addition to interviews to collect similar information from patients [148, 162].

3.4.8. Prototype Testing

Many included studies were non-experimental or pilot studies used to assess the feasibility and measure usability. These formative studies can be considered a step for development before releasing and testing a mature chatbot in a RCT. For example, one study using a chatbot for an exercise intervention organized a three-week formative usability study [163] to identify issues and make revisions before conducting an RCT [155].

3.4.9. Wizard of Oz (WoZ) Procedure

One study utilized the "Wizard of Oz" (WoZ) procedure [157] (where the technology is controlled by a human interface in chatbot development) as a step in their chatbot development. This procedure is administered by engaging participants in a 20-minute conversation with a simulated chatbot that was not automated but controlled manually by a researcher answering questions on the back end [157]. This step was developed to understand how the chatbot should interact with humans in a natural setting and to collect content-related information directly from participants [157].

3.4.10. Use of Existing Literature to Gain Evidence-Based Knowledge for Development.

In ten studies, initial literature reviews were completed to gain evidence-based knowledge to guide chatbot development [148, 151, 152, 155, 158-163]. In three of these ten studies, a literature review was used to develop content from evidence-based sources, including self-management practices, clinical guidelines, and systematic reviews [148, 152, 158]. A mental wellness study incorporated this step into development by reviewing the psychological theories and practices used to create the lessons the chatbot would deliver [151]. In another study, a literature review of existing health behavior change models was conducted to understand the pros and cons of each model, and to guide the development of a novel behavior change model to structure the chatbot's content [162]. In one study, grey literature was sourced through online chat threads for a youth helpline, so researchers could better understand content topic preferences and expected answers [160]. Finally, two of these ten studies reviewed the literature to learn more about reward systems and to identify a theoretical basis for chatbot development [155, 163].

3.4.11. Patient Engagement and Public Involvement

Overall, the reporting of patient engagement in our included studies was limited making an assessment of PPI using the GRIPP2 challenging. Though eight studies in our review reported involving patients, five provided inadequate detail, making assessing patient involvement impossible [148, 149, 151, 162, 163]. Specifically, these studies did not report on the aim of PPI, did not clearly articulate their methods, or did not discuss the role of PPI in their outcomes. The remaining eight studies were not evaluated using the GRIPP2 because they did not report development approaches at all [153, 154, 156] or did not involve patients in the reported approaches [152, 155, 158, 159, 161].

Of the three studies we assessed using the GRIPP2, one study scored three points on the GRIPP2 Field [38], with the other two scoring one point [150, 160]. Figueroa et al.'s study scored 3/5 on the GRIPP2 scale [157]. This study provided a clear description of the methods used for PPI, commenting on how PPI influenced the study and on successful and unsuccessful aspects of the study relating to PPI [157]. This study was also the only one that described four different approaches used for development, including co-design workshops, interviews, WoZ, and prototype testing. The authors noted that their co-design sessions "brought unexpected participant preferences and wishes, which were useful in developing subsequent versions" of their chatbot [157]. Further, they recognized the importance of engaging patients in design, testing, and dissemination to develop chatbot interventions that participants would use and benefit from. The remaining two studies, one by Gabrielli et al. and the other by Maenhout et al., were each awarded a single point on the GRIPP2 for clearly describing the methods used for PPI [150, 160]. The reporting was such that future researchers could replicate similar development approaches to actively engage patients in research design.

3.5. Discussion

In this review, we described the non-technical approaches taken for chatbot development and evaluated the extent of patient engagement using the GRIPP2. While promising approaches were shared about the non-technical steps associated with chatbot development, the level of detail provided was often low, including how patients were involved in the process.

The limited level of detail speaks to the need to prioritize frameworks for implementing digital health tools [164, 165]. This will involve a focus towards increased formative, development, and feasibility studies and a shift to implementation research that considers embedding and sustaining interventions in context [164, 165]. A more detailed focus on the developmental stages and implementation process in research would allow increased replicability of developmental approaches that actively engage patients and progress the field of chatbot research from the end-user perspective. An example of this focus on the implementation process includes the formative work conducted by Islam and Chaudary while developing a chatbot to support the healthcare needs of patients during the recent COVID-19 pandemic [166]. Their work is an example of detailed documentation of a replicable multi-phased chatbot design study, offering guidance for future research in this area [166]. Additional focus on implementation will ensure the production and monitoring of chatbots that provide quality care and service to patients across short- and long-term timelines [164]. This strategic planning also holds promise to better respond to the requirements of diverse user cohorts, especially those with lower levels of digital health literacy [167].

While an attempt was made to evaluate the extent of the patient engagement process by the GRIPP2 patient engagement checklist, due to limited detail of reporting, this was only possible in three studies [150, 157, 160]. Many digital health solutions are plagued with low uptake and poor usability as they were developed with minimal patient involvement [120]. As user-centered design and patient engagement are known to improve the quality of research, utilizing engagement approaches throughout the research continuum could result in the identification of system requirements that would be otherwise missed, as well as result in a better understanding of patient needs, higher intervention engagement, and increased intervention effectiveness [56]. Some of the approaches we have identified in this review, including co-design workshops, the WoZ approach, patient interviews, and iterative prototype testing, represent ways researchers can creatively and actively engage patients throughout the development process. Co-design workshops foster a richer understanding of what patients "know, feel, and even dream" [168]. The WoZ approach is a widely accepted evaluation and prototyping methodology for developing human-computer interaction technology [169]. Engaging patients in iterative prototyping and user testing cycles has proven to improve the ease of use and adoption of these interventions [170]. In alignment with the literature, we recommend that researchers taking on health chatbot development projects consider adopting approaches such as co-design workshops, interviews, WoZ, and prototype testing.

Despite the available evidence supporting the benefit of patient involvement in intervention development, there are reasons why approaches that do not directly or actively involve patients may be more appealing to researchers. This notably includes challenges associated with recruitment, particularly when trying to avoid recruitment bias, and the time and resource intensity associated with the overall process [141]. The scarcity of patient involvement may also be related to an underappreciation of the potential benefits of patient involvement in digital health research and a limited understanding of how best to get patients involved [141].

Researchers and practitioners should be aware that there are many different approaches, strategies, and models to engaging patients in chatbot development. We have summarized some approaches in this review, and resources such as the Strategy for Patient-Oriented Research patient engagement framework and the patient engagement in research plan offer practical information to guide patient involvement in the development process [127, 171]. Patients can participate at all stages, helping to define healthcare problems, identify solutions, participate as co-designers of an intervention and refine the evaluation process [140]. **Figure 3.4** offers direction in informing future research in patient-oriented chatbot development for lifestyle and wellness interventions, including the application of multifaceted means of patient engagement, use and thorough documentation of approaches to enhance chatbot development, and clear and replicable reporting of the formative stages of development.

Figure 3.4. Informing areas of future research in patient-oriented chatbot development for lifestyle and wellness interventions.

3.5.1 Strengths

We searched nine of the most relevant bibliographic databases for medical and technology research for this review. No restrictions were placed on the year of publication, country of publication, journal, or study setting. Our study team consisted of multidisciplinary research and healthcare professionals with relevant expertise who provided direction at each review phase. This review was guided using an established framework proposed by Arksey and O'Malley [142].

3.5.2 Limitations

This review focused on simple voice- or text-based chatbots that engaged in two-way communication with human users. This led to the exclusion of other forms of conversational agent technology (i.e. embodied conversation agents, humanoid/social robots, wearable technology, IoT, virtual avatars, interactive voice assistants, etc.) that may have resulted in the finding of additional development and engagement approaches that were not covered in our review. Our review excluded literature from conference proceedings, protocol papers, and other papers lacking an intervention. Moreover, although our proportionate agreement was 0.967 at the title and abstract screening stage, there was only "fair" agreement between reviewers (Cohen's Kappa = 0.309). This "fair" agreement between researchers highlights the challenges in reviewing a heterogeneous body of literature. With ongoing meetings and refinement of our inclusion/exclusion criteria, the Cohen's Kappa statistic improved to an "almost perfect" agreement at the full-text review stage (Cohen's Kappa = 0.843). Additionally, due to the limited detail available within the included studies, our team could not conclusively assess patient involvement in chatbot development; greater attention to reporting patient involvement in chatbot development and testing in future research will help with this limitation. Finally, we acknowledge that scoping reviews have numerous shortcomings, including limitations of rigour and potential bias stemming from the absence of a quality assessment, among others [172]. However, the literature on chatbot technology remains highly heterogeneous at this time, and scoping review provided a systematic method to map the current state of the literature.

3.6. Conclusion

In conclusion, this review provides a menu of options that can be used for the non-technical steps associated with chatbot development for interventions supporting lifestyle and wellness interventions. The identified study limitations hold promise to guide the inclusion of patient engagement and the improved documentation of the engagement and development of chatbots in future healthcare interventions. Given the importance of end-user involvement in the development of digital technology, it is our hope that future research on chatbot development will take the opportunity to carry out a more systematic reporting of the chatbot development and implementation process and will actively engage patients as key members of the co-development process.

CHAPTER 4: FREQUENTLY ASKED QUESTION (FAQ) ANSWERING CHATBOT FOR ONLINE MENTAL WELLNESS PROGRAM: DEVELOPMENT AND PILOT EVALUATION

Authors: Chikku Sadasivan, Ashley Hyde, Emily Johnson, Eleni Stroulia, Puneeta Tandon

4.1. ABSTRACT

Introduction

Chatbots are computer programs that interact with humans through natural language conversations. A frequently asked question (FAQ) answering chatbot is a program designed to answer commonly asked questions. Chatbots can provide users with convenient, 24/7, personalized support and show promise in healthcare to support people living with chronic diseases. While chatbots, like other digital health interventions, are promising tools for chronic disease management, they face challenges such as limited user adoption and low engagement. One way to address these challenges is to involve patients in development through user-centred design.

Objective

In this pilot study, we conducted a formative evaluation of an online, frequently asked questions (FAQ) answering chatbot. This chatbot was deployed on a 12-week online mental wellness program for patients with chronic diseases run by our research team. This paper describes the non-technical aspects of the chatbot's development, approaches to involving patients in development, and pilot evaluation results.

Methods

A FAQ-answering chatbot ("Liv") was developed through a multi-phased, multi-method design using a qualitative descriptive approach. Liv was then deployed on an online mental wellness program and evaluated for acceptability, usability, and user engagement.

Results

Liv was deployed for 120 days on our online wellness program. During the deployment, there were 259 instances where a unique user sent a message to Liv and received a response, with 175 instances of active engagement (back-and-forth user-chatbot interaction). The engagement was highest during the first 30 days of deployment. The chatbot had a resolution rate of 33%. Results from the chatbot usability questionnaire included an overall score of 50.8 (below average) for usability, with successes in specific areas, including its navigation and ease of use. Findings from qualitative interviews included comments on Liv's personality, the convenience of knowing the chatbot would escalate, and experiences with Liv and her improvement over time.

Conclusions

FAQ chatbots may be an engaging way to provide patient support in online mental wellness programming. Including patients in development may improve the chatbots' acceptability, usability, and user engagement. Future work is warranted that prioritizes patient engagement and builds upon this pilot evaluation's data to create a more usable chatbot.

Keywords: chatbot, digital health, virtual assistant, chronic disease, online mental wellness programming, patient involvement, participatory design

4.2. Introduction

Chatbots are computer programs that use artificial intelligence to interact with humans through natural language conversations in text or speech [1]. Today, messaging applications are a routine means of communication, with over 6 billion people using SMS-capable mobile phones [65]. Along with this increase in the use of messaging applications, chatbots have also been on the rise in areas such as customer service, sales, personal assistance, healthcare, education, banking and finance, human resources, entertainment, and travel [6, 7, 47-49]. It is anticipated that many people living with chronic disease - one in three Canadians [33] - have also encountered chatbots through at least one of these means. Chatbots can offer interventions directly or in conjunction with interventions delivered through web-based or smartphone modalities.

Like any other website users, patients have questions and need answers. Traditionally, FAQ webpages were widely used to answer questions but can be frustrating for website users due to information overload [173]. For example, website users may have to read through many different questions before finding the one that matches their concern, or the wording and terminology on the FAQ page may not be something the user is comfortable with [173]. In a study by Chatterjee et al. that described the current limitations of FAQ webpages, it was reported that when questions are not answered, website users will bypass FAQ webpages and use other routes to have their questions answered, including by email, phone call, or community answering forums [174]. Responding to user inquiries only via emails and phone calls can result in long wait times to receive support, reduce customer satisfaction, and be intensive on resources and staff [174]. To improve Q&A service, FAQ chatbots can be implemented as a potential solution [173, 175].

A frequently asked questions (FAQ) chatbot is designed to answer common questions people have [176]. While chatbots, like FAQ webpages, cannot answer all patient inquiries, they aim to provide more human-like, personal, and immediate responses through natural language interactions [175]. Many FAQ chatbots showed promise when developed in response to the recent COVID-19 pandemic to provide an efficient source of information [177-179]. These chatbots were developed quickly and served to efficiently disseminate information to the public about the ongoing pandemic while provider resources were limited [177-179]. For example, the WHO created a simple chatbot deployed on WhatsApp to answer questions and communicate information about COVID-19 risks globally [178]. FAQ chatbots were also used to screen positive cases and answer questions about COVID-19 [179]. Post-pandemic, researchers like Bharti et al. have sought to use FAQ chatbots to provide patient education and increase health literacy for patients facing healthcare accessibility barriers [180]. Bharti et al. found that their chatbot increased access to healthcare information, reliably detected various common diseases, and suggested treatment remedies in an engaging and personalized manner [180].

Over the past five years, our research group has been involved in developing and evaluating a 12-week online mental wellness program to support individuals living with chronic disease [57, 181]. This wellness program includes mindful movement (yoga, tai chi, and chair exercise), energizing breathwork practices, guided meditation, and positive psychology content [57]. Previous programming has positively impacted patient stress, mental health, and health-related quality of life (HRQOL) [57, 181]. Traditionally, through these studies, patient questions were answered over email. This was resource intensive for study staff and resulted in delayed response time depending on staff availability. Our latest version of the program, EMPOWER, contains similar mindful movement, breathwork, meditation, and psychology content as previous programs and is currently being trialed for patients with various chronic diseases such as (heart failure, chronic kidney disease, primary biliary cholangitis, etc.) [182]. Given the potential benefits of a FAQ chatbot for the newest program, EMPOWER, my project aimed to experiment with developing and deploying this chatbot as part of the RCT.

Although our team realized the numerous benefits of including a chatbot, such as increased efficiency for the study team and improved patient support, we also learned that potential challenges of chatbot implementation, like limited user adoption and low engagement, have been brought up commonly in the literature [41-44]. This is concerning as engagement is essential to intervention efficacy [42]. Experts suggest that these challenges can be mitigated through user-centred design (UCD) principles and by including patients in the development process [53, 54, 120]. Reported benefits include creating technology that is more relevant, usable, and effective [53]. Given our work showing limited patient involvement in the chatbot design process [76], our goals were to: (1) develop a FAQ chatbot ("Liv") through a multi-phased, multi-method design using a qualitative descriptive approach and (2) deploy the chatbot on an online mental wellness program and evaluate for acceptability, usability, and user engagement.

4.3. Methods

Recognizing the often-cyclical nature of software development, we used a multi-phased approach (**Figure 4.1**). The multi-phased approach was guided by steps from the Canadian

Institutes of Health Research (CIHR) Strategy for Patient-Oriented Research (SPOR) framework [127] and the Chatbot Development Life Cycle [70]. The CIHR SPOR patient engagement framework guides researchers in meaningfully engaging patients throughout the research process, including identifying problems and implementing solutions [127]. This framework and the Chatbot Development Life Cycle ensured the patient's experiential knowledge was central to our multi-phased design [127]. The Chatbot Development Life Cycle is a framework that covers the chatbot building process with 11 steps repeated in an iterative development process [70]. Development phases were mapped onto the chatbot development life cycle to identify the cycle stages in which patients were involved (**Figure 4.2**).

Figure 4.1. Multi-phased chatbot design timeline.

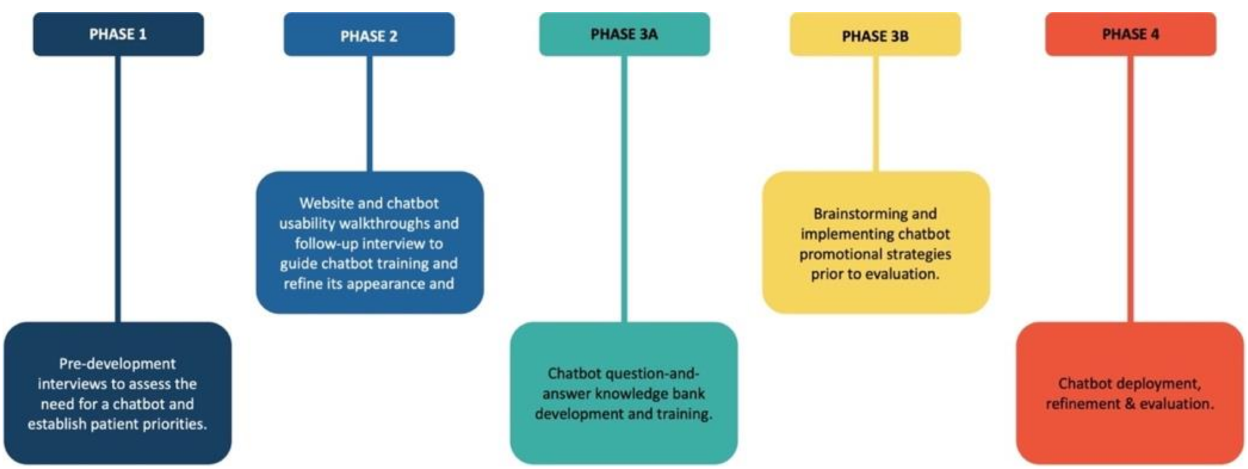
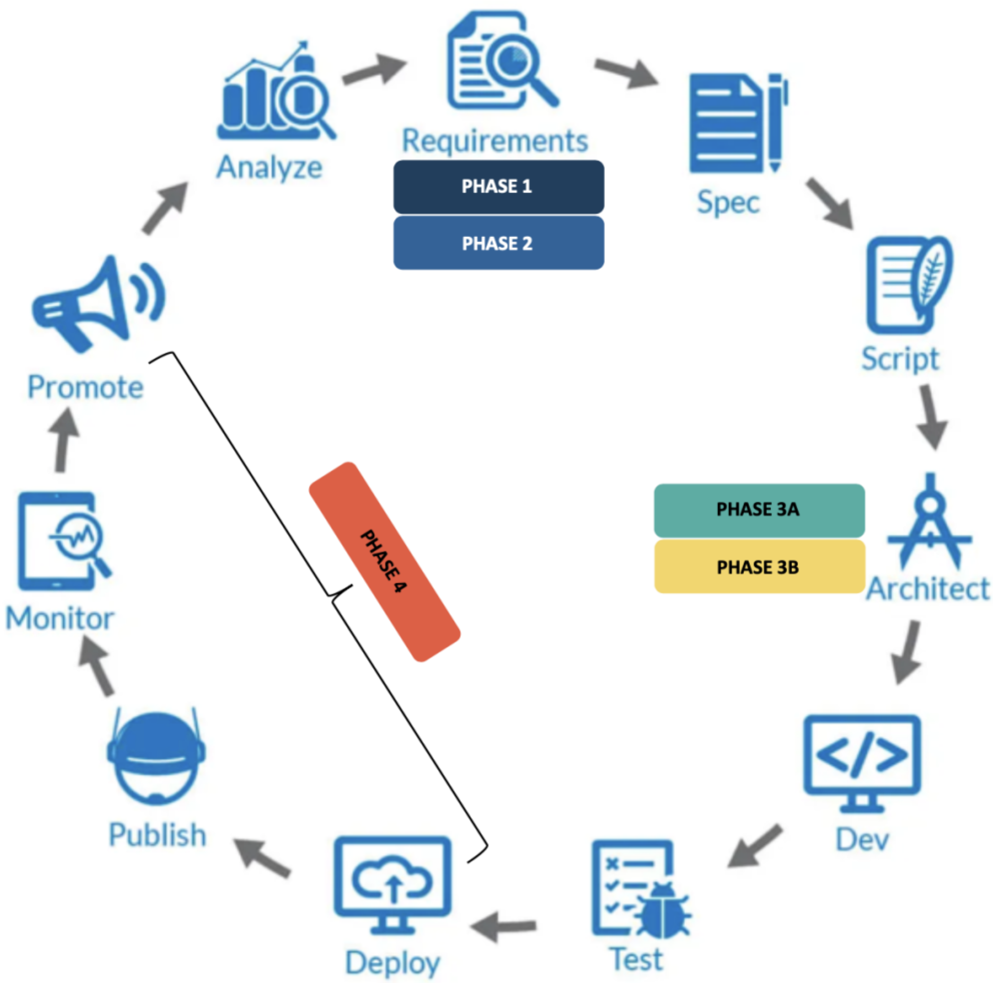


Figure 4.2. Multi-phased chatbot design mapped onto chatbot development life cycle.



To understand patient needs and perspectives, multiple methods (qualitative and quantitative) were used throughout the development and evaluation phases. A qualitative descriptive approach was used to establish patient priorities and utilize patient input to guide development and iteration. Data were collected via semi-structured interviews [183], with interviews conducted by EMPOWER staff via Zoom at a time convenient for participants. Interviews were recorded and transcribed verbatim. Analysis followed a thematic approach whereby interviews were coded, with codes combined into larger categories and themes [184-186]. Participants were recruited from a sample of patients who participated in a previous version of the 12-week online mental wellness program (Peace Power Pack) until data saturation [187]. Participants were purposively sampled [184], and invitations were sent via email to include higher and lower technical proficiency users and users who were both experienced and inexperienced with the program.

Quantitative approaches were used to evaluate the chatbot's usability and user engagement in phase 4. Usability was assessed by using the chatbot usability questionnaire (CUQ) [188], a validated questionnaire designed to evaluate a chatbot's personality, ease-of-use, intent recognition, response clarity, and error management [188]. The CUQ consists of sixteen balanced questions to assess chatbot usability, with eight questions relating to positive and eight about negative chatbot aspects [188]. Questions are scored on a five-point Likert scale with an overall score calculated out of 100 [188]. All CUQ data were analyzed in Microsoft Excel to calculate descriptive statistics. Additionally, engagement was measured through Zendesk Explore data collected throughout the chatbot's deployment period [189]. Engagement data included the number of messages sent, conversations with active engagement (back-and-forth user-chatbot interaction), chatbot resolution rate, escalations, and engagement patterns across time. Conversation log data was analyzed to understand chatbot usage patterns and engagement with the chatbot over time and to collect frequency data on specific question topics. Finally, information on participant technical proficiency was compiled based on self-report data and estimated by patients using a scale used in our previous studies [190]. Categories for this technical proficiency scale included: 1) technology expert: can do everything wanted/needed using technology and can successfully troubleshoot, 2) highly: can do almost everything wanted/needed using technology and can usually figure out new tasks, 3) moderately: can do on average most tasks but can find it challenging to figure things out beyond comfort level, 4) sort of: can do some tasks but there are a lot of areas that are uncomfortable, 5) a little: can interact somewhat with technology but uses devices minimally.

4.3.1. Phase 1: Pre-development interviews to assess the need for a chatbot and establish patient priorities.

The first phase (May 4, 2022 to May 26, 2022) was conducted to identify the need for developing a chatbot, understand patient perspectives on using chatbots, and establish patient priorities to better understand requirements that would need to be met for the chatbot. Interviews were conducted with seven participants. For this phase, higher and lower technical proficiency users were recruited from our previous Peace Power Pack (PPP) mental wellness program [57, 181]. Technical proficiency was defined by the frequency of required technical support during the previous program. Higher proficiency users either rarely or never sent emails about unresolved issues. Lower proficiency users frequently experienced issues and sent multiple emails for support, without being able to resolve these issues on their own.

Table 4.1. Semi-structured interview guide for phase 1.

1. Can you tell me about what you understand a chatbot to be?
2. We're exploring making improvements to the app through integration of a chatbot. What conversations would you feel comfortable having with the chatbot? What topics would you want a chatbot to answer for you?
3. When you think back to when you first started the program, what were some of the questions you had?
4. While using the website, where did you encounter problems or when was the website most troublesome to use?

4.3.2. Phase 2: Website and chatbot usability walkthroughs and follow-up interviews to guide chatbot training and refine its appearance and feel.

The second phase (December 2022 to January 2023) aimed to evaluate the usability of the prototype chatbot and web platform, identify areas to refine for chatbot training, and gather feedback on the chatbot's content, appearance, and overall user experience. During this phase, the prototype chatbot was still untrained but was embedded in the web platform for preliminary usability testing. One-to-one website usability walkthroughs were conducted, during which participants were asked to complete six navigation tasks. These included logging on to the website, locating a video activity, performing an interactive activity, navigating to the website's leaderboard, and communicating with the chatbot. Following the walkthroughs, interviews were conducted with eight participants. For this phase, both experienced and inexperienced users (with our platform) were recruited.

Table 4.2. Semi-structured interview guide for web platform and chatbot usability walkthroughs

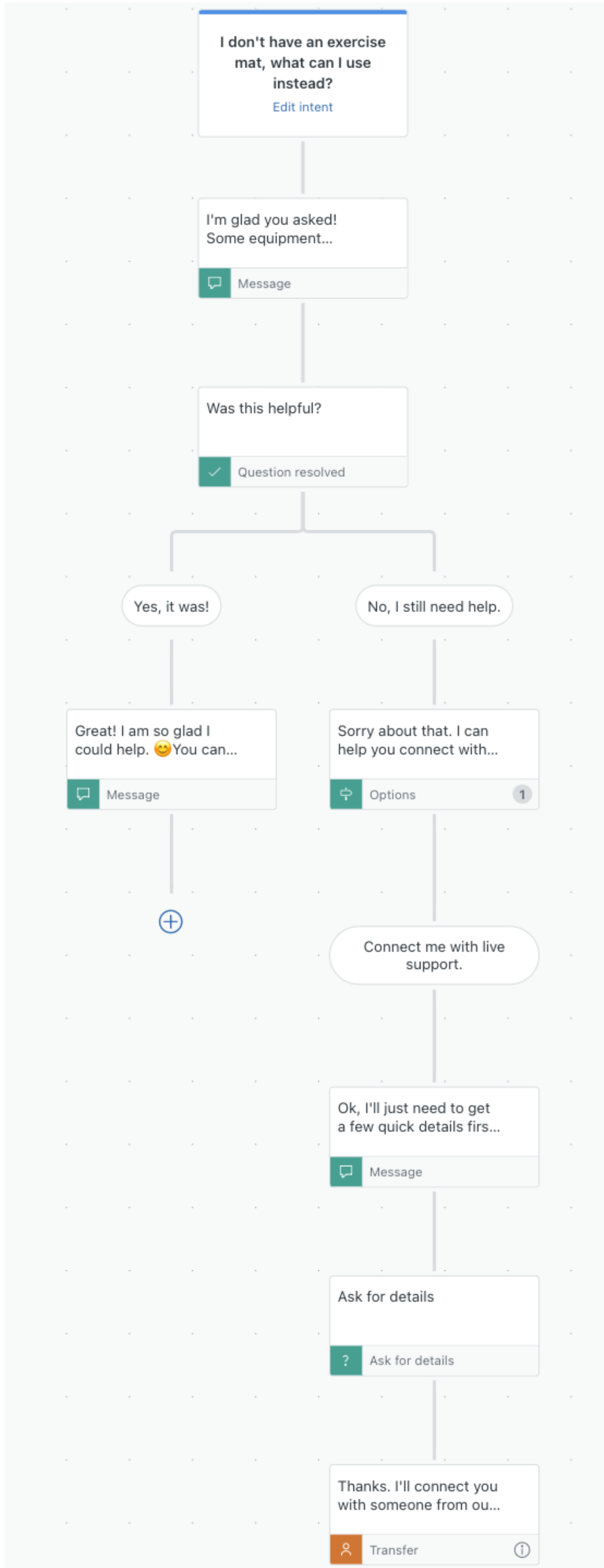
1. In the context of this program, what do you think the virtual assistant could do for you?
2. Can you tell me about your experience with navigating through the website tasks today? How could a chatbot help?
3. What was your experience with finding and using our virtual assistant?
4. How could we increase awareness about this virtual assistant to get people to use it? We're thinking of an information video on the help page describing what it does and how our team members will be in the loop, what else could we do?

4.3.3. Phase 3A: Chatbot question-and-answer knowledge bank development and training.

The question-and-answer knowledge bank was developed by incorporating results from phases 1 and 2, past program email records, evidence-based information sources for program content, and collaboration with team knowledge experts (physician, psychologist, dietitian, mindfulness expert, exercise physiologist, etc.). The questions and answers for the chatbot were organized in a spreadsheet to cover topics like exercise, nutrition, wellness, technological assistance, motivation, and small talk. To create a more human-like and engaging experience, the chatbot's personality was crafted in collaboration with the psychologist who contributed to developing the wellness intervention. Collaboration with our program psychologist to aimed to make the chatbot's initial responses to inquiries validating, empathetic, and motivating for users. For example, to be validating Liv's initial response to an inquiry would state: "That's a great question that I get quite often!" Finally, the chatbot was named Liv, a greeting message was established, and its web widget appearance was customized to align with the overall design of the web platform.

The knowledge bank was used to develop the chatbot using Zendesk's answer bot flow builder (**Figure 4.3**) [191]. Questions were inputted as intents, and the answer flow was constructed in a decision tree format that allowed for conversation escalation to a human agent [192]. Chatbot fallback, a message triggered when a chatbot does not recognize a user's intent and offers the opportunity to escalate the conversation to a human [193], was set up and tested to allow for integration with Zendesk's agent dashboard for live support responses. The Zendesk agent dashboard enabled the chatbot's fallback to a human agent for correct identification of user intents, to offer timely patient support, and to iteratively refine the chatbot's training during phase 4. Finally, question-answer flows were proofread and previewed to test for functionality using Zendesk's conversation preview mode.

Figure 4.3. Example question in Zendesk's answer bot flow builder.


I don't have an exercise mat, what can I use instead?
Edit intent
I'm glad you asked! Some equipment...
Message
Was this helpful?
Question resolved
Yes, it was!
No, I still need help.
Great! I am so glad I could help. 😊You can...
Message
Sorry about that. I can help you connect with...
Options
1
Connect me with live support.
Ok, I'll just need to get a few quick details firs...
Message
Ask for details
Ask for details
Thanks. I'll connect you with someone from ou...
Transfer

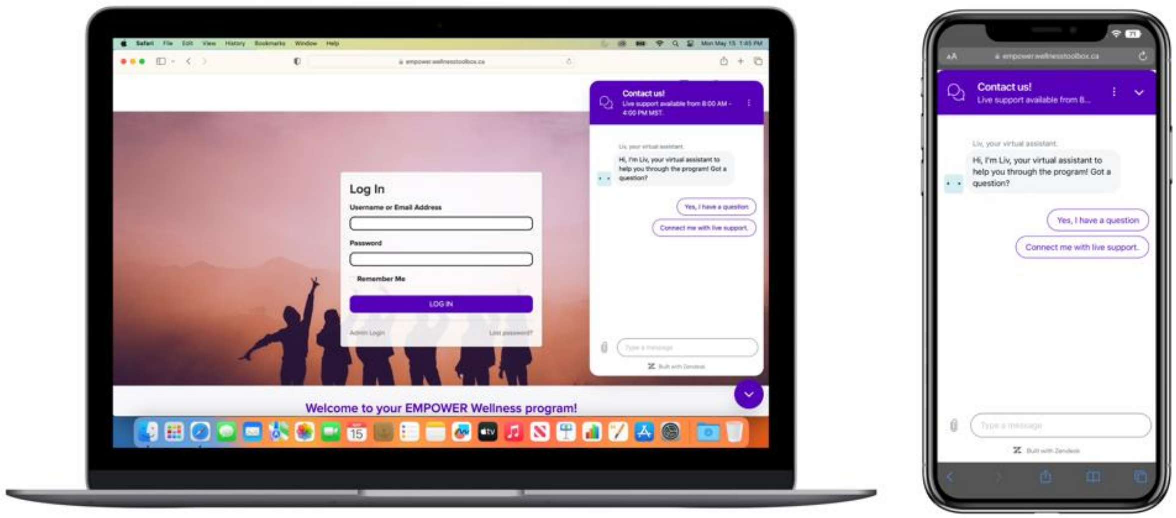
4.3.4. Phase 3B: Brainstorming and implementing chatbot promotional strategies prior to evaluation.

This phase was conducted in concert with the final stages of development of the digital wellness program to maximize engagement with the chatbot on the web platform. Based on the feedback obtained from participants in phase 2 interviews, the chatbot was embedded on all web pages of the platform and chatbot support reminders were scheduled to send out to program participants via a weekly email newsletter. Additionally, a promotional video was created to explain the chatbot's features, capabilities, and guidelines for its use [194]. This video was included in the "before you begin" program onboarding content and on the website's support page. Finally, information was provided to patients through Zoom virtual onboarding sessions to guide them on accessing the chatbot for support.

4.3.5. Phase 4: Chatbot deployment, refinement & evaluation

During Phase 4, the chatbot (Liv) was deployed on the EMPOWER program via a chat widget at the corner of the web page (**Figure 4.4**) from February 12, 2023, to June 12, 2023. Liv could be used on laptops, desktop computers, and mobile devices. EMPOWER program staff supported the chatbot on Zendesk's agent dashboard to assist Liv from the back end whenever patients' inquiries were not resolved and patients opted to escalate the conversation. Liv was monitored and tested for functionality throughout the deployment period. Questions that consistently came into live support were added to Liv's knowledge bank. Additionally, user intents that had already been trained but were incorrectly identified by Liv were refined in the knowledge bank.

Figure 4.4. Liv's deployment on the EMPOWER web platform.



Several strategies were used to evaluate Liv's usability, acceptability, and engagement. Usability was evaluated in the fourth week of the program using the CUQ [188]. Acceptability was evaluated by gathering impressions and experiences of patients' chatbot usage through qualitative interviews **(Table 4.3)**. Participants were purposively sampled based on self-reported technical proficiency, program experience, and age. Engagement was measured through Zendesk Explore data collected throughout the chatbot's deployment period [189]. Engagement data included the number of messages sent, conversations with active engagement (back-and-forth user-chatbot interaction), chatbot resolution rate, escalations, and engagement patterns across time. Finally, conversation log analysis of chatbot and live support chat histories was completed to understand the conversations participants had to guide future development.

Table 4.3. Semi-structured interview guide to evaluate Liv's acceptability, usability, and user engagement.

1. Can you tell me about your experiences with our virtual assistant, Liv?
2. If Liv worked for you, what did you like about her?
3. If Liv didn't work for you, what did you think?
4. What was it like to connect with our live support team?
5. What are your thoughts on the ease of finding and using the virtual assistant?
6. Please share your feelings about Liv's personality and conversational style.
7. How did Liv compare to other virtual assistants you've interacted with?

4.4. Results

4.4.1. Phase 1: qualitative findings from pre-development patient interviews to assess the need for a chatbot and establish patient priorities.

Seven participants were included in phase 1 interviews to assess the need for a chatbot and critical requirements **(Table 4.4)**. Of these participants, three were more technologically proficient users who either rarely or never sent emails about unresolved issues during the PPP program (ID: 1, 2, 3), and four were less technologically proficient users who frequently experienced issues and sent multiple emails for support, without being able to resolve these issues on their own (ID: 4, 5, 6, 7). All participants were female, with an age range of 62 – 74.

Table 4.4. Phase 1 interview participant demographics.

ID	Age	Sex	Education Level
PPP1	72	F	University degree
PPP2	68	F	Non-university certificate/diploma
PPP3	74	F	University degree
PPP4	65	F	University degree
PPP5	73	F	Non-university certificate/diploma
PPP6	62	F	No post-secondary
PPP7	65	F	University degree

Phase 1 interviews identified several themes about participants' past experiences with chatbots, a "give it a try" mindset on using the chatbot, identifying the need to escalate to a human, and clarifying the purpose of a chatbot. The identified themes, descriptions, and sample interview quotes are summarized below (**Table 4.5**).

Table 4.5. Phase 1 interview thematic findings and sample interview quotes.

Theme	Description of theme	Sample interview quotes
Past experiences with chatbots led to hesitancy	Participants shared their perceptions of chatbots, largely colored by negative experiences with bots on other platforms.	"I would ask it a question, a simple question. And it would come back with an answer that was completely unrelated. I've never had a good experience with a chatbot." (PPP4)
		"I find that quite often, that chatbots are not all that great with the technical problems." (PPP1)
		"When you're left kind of high and dry, it's annoying." (PPP1)
Adopting a "Give it a try" mindset	Though several participants were influenced by their past experiences with chatbots, they described being willing to engage with a chatbot on our wellness platform. They did, however, acknowledge that the chatbot should be able to escalate their issue promptly should it not be able to answer their questions.	"Two or three times giving [the chatbot] a try see what it was and see if it met my needs." (PPP1)
		"If the question is "how do I use this? Or where do I find this? Maybe they're useful. So I could see, instead of emailing a person and interrupting their work, if that was a simple question, then I could see there'd be some value." (PPP2)
		"If the chatbot couldn't answer my question, then it should relay that to you, and then you get back to me." (PPP6)
		"If it can't answer the question, say, who to contact, email or a phone number." (PPP1)
Clarifying the purpose of the chatbot	Participants described different interpretations of what a chatbot was, often confusing it with chat boxes, forums, or chat rooms. Further, they shared different expectations of what they felt the chatbot should be able to achieve.	"For me it's like going in and speaking with other people who are experiencing similar issues or maybe a discussion on what other people are going through, and maybe I've had the same experience and you exchange ideas." (PPP3)
		"It's like something that participants can chat back and forth to one another." (PPP6)
		"It's something that runs a chat room. But, that's the only thing I can think of. It's a place to chat." (PPP1)

4.4.2. Phase 2: Qualitative findings from website and chatbot usability walkthroughs and follow-up interview to guide chatbot training and refine its appearance and feel.

Eight participants were included in phase 2 interviews to examine website and chatbot usability for chatbot training and refinement (**Table 4.6**). Of these participants, six were users who participated in a previous PPP program (ID: 1, 2, 3, 4, 5, 6) and two had no prior experience with the PPP program (ID: 7, 8). All but one of the participants were female and the age range was 62 – 74.

Table 4.6. Phase 2 interview participant demographics

ID	Age	Sex	Education Level
UW1	72	F	University degree
UW2	74	F	University degree
UW3	62	F	No post-secondary
UW4	68	F	Non-university certificate/diploma
UW5	67	F	University degree
UW6	65	F	University degree
UW7	66	F	University degree
UW8	67	M	University degree

Phase 2 interviews led to feedback on the chatbot and website, including overall perceptions and reactions to the chatbot, suggestions for improvement, and highlighting the need for chatbot promotion. The identified themes, descriptions, and sample interview quotes are summarized below (**Table 4.7**).

Table 4.7. Phase 2 interview thematic findings and sample interview quotes.

Theme	Description of theme	Sample interview quotes
Initial chatbot impressions	Participants shared largely positive impressions of the chatbot prototype, appreciating the promptness of the chatbot response and its efficiency relative to emails. Several participants described the strong influence of their past experiences with chatbots and how this might influence their engagement with our chatbot.	"I was sending emails left, right and center for things that I couldn't get done. So I'm going to enjoy having my simple questions or less complicated answered by the chat, and it's really nice that the chat will then go off to someone else who could answer the more difficult questions." (UW3) "Before chatbots existed, all you could do is ask a question by email, and then god knows how long it would take for you to get an answer. I ended up enjoying the chatbot because I would get an almost instant response." (UW6) One participant noted that "it's an almost an instant reply, which is good, because sometimes I know I've been on virtual assistants, and I've had to wait a few

		minutes, sometimes even an hour. So that can be frustrating for sure" and suggested that a response "within one or two minutes" was reasonable. (UW2) "I think, that it's just like every other one I've ever used. So that's a familiarity that makes it comfortable." (UW8) "I try to avoid them; I like talking to live people." (UW5)
Identifying areas of improvement	Participants identified areas for improvement of our chatbot prototype. These include the location of the chatbot widget on the website and its ability to help with navigation tasks like locating resources on the site.	Finding the chatbot "was more challenging than expected because everything else is kind of on the left-hand side, and then that was up in the right hand top corner. I wasn't looking for it there." (UW5) "The little chat thing is at the bottom and I'm wondering if everybody would catch that" (UW1). Another participant noted that "I think having it in a prominent location, front and center of your main landing page - I think that's important." (UW8) "It's just at the beginning. I had a bit of a problem there. I think if you recall finding the nutrition and the menu that one particular recipe, the Greek wrap." (UW7)
Highlighting the need for promotion	Participants emphasized the need to promote the chatbot to new website users, including emphasizing what types of issues the chatbot can solve. They suggested using a chatbot tutorial and reminding patients that the chatbot is there to support them.	"You have to know the limitations and parameters of what you" can ask the virtual assistant, so this participant suggested to create "an outline that you would want to tell people what it was capable of doing." (UW4) Another participant described this same ideas as "a little tutorial on the chat with our team." (UW3) "Sometimes people forget what they have, what resources they have. If you had something that even popped up on the screen. I'm not sure when that would be, and just said, hey, remember, I'm here to help you if you need me kind of thing." (UW4) "There should be something right at the very start to say you know, don't hesitate to use the chat, help chat area." (UW1)

4.4.3. Phase 4: Findings from chatbot deployment, refinement & evaluation

Liv was deployed 120 days from February 12, 2023, to June 12, 2023. During deployment, 130 participants participated in the EMPOWER, a 12-week online wellness program. Of the 130 participants, 45 (96% female, age 51 – 81 years) reported using the chatbot during the fourth week of the program and filled out the CUQ (**Table 4.8**). Additionally, 20 participants (100%

female, age ranging from 51 – 76 years) took part in end-of-program interviews where they commented on their experiences with Liv (**Table 4.9**).

Table 4.8. Characteristics of survey respondents.

Age Group	Self-Reported Technical Proficiency				
	Expert	Highly	Moderately	Sort of	Total
50-59	4 (9%)	3 (7%)	1 (2%)	1 (2%)	9 (20%)
60-69	4 (9%)	10 (22%)	11 (25%)	1 (2%)	26 (58%)
70+	4 (9%)	2 (4%)	3 (7%)	1 (2%)	10 (22%)
Total	12 (27%)	15 (33%)	15 (33%)	3 (6%)	45 (100%)

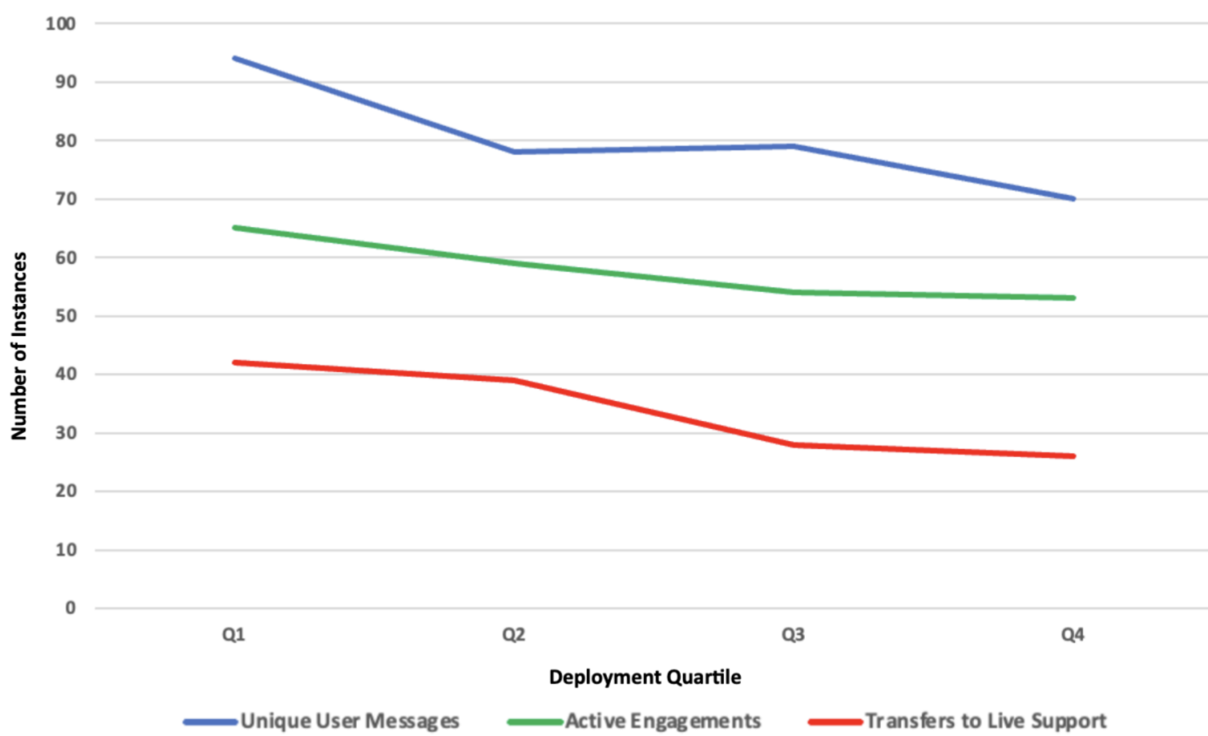
Table 4.9. Phase 4 interview participant demographics

ID	Age	Education Level	Self-Reported Technical Proficiency
11	72	College/University Degree	Moderately
23	61	College/University Degree	Moderately
25	68	High School Diploma	Moderately
34	62	College/University Degree	Highly
35	65	College/University Degree	Moderately
44	72	College/University Degree	Highly
52	63	Some College/University	Moderately
53	61	College/University Degree	Moderately
59	56	College/University Degree	Highly
60	65	Some College/University	Highly
62	66	College/University Degree	Expert
63	64	Non-university Certificate/Diploma	Moderately
75	71	College/University Degree	Expert
78	60	Some College/University	Moderately
81	66	High School Diploma	Highly
84	62	Some College/University	Expert
95	55	College/University Degree	Moderately
109	51	College/University Degree	Expert
124	63	College/University Degree	Highly
129	76	High School Diploma	Moderately

<u>Chatbot Engagement, Performance, and Usage</u>

During the deployment period, there were 259 instances where a unique user sent a message to Liv and received a response. 175 of these 259 instances involved more active engagement with multiple messages back and forth between users and Liv. There were 201 responses to Liv’s resolution feedback prompt, with 67/201 (33%) resolved and 134/201 (67%) unresolved. 131 participants escalated their conversations to human agents at live support to resolve their concerns. The number of unique user messages, active engagements, and transfers to live support over four quartiles of the 120-day deployment period are compiled below (**Figure 4.5).** Each quartile represents a 30-day period. Engagement via messages sent to Liv and transfers to live support were highest during the first 30 days of deployment.

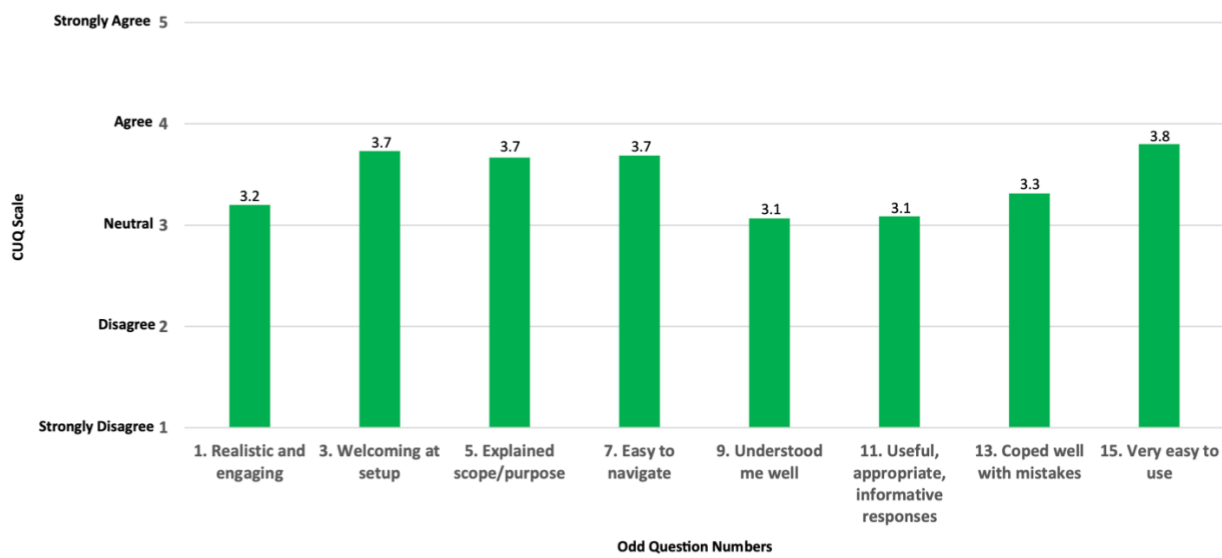
Figure 4.5. Engagement data across the 120-day deployment period.



Chatbot Usability

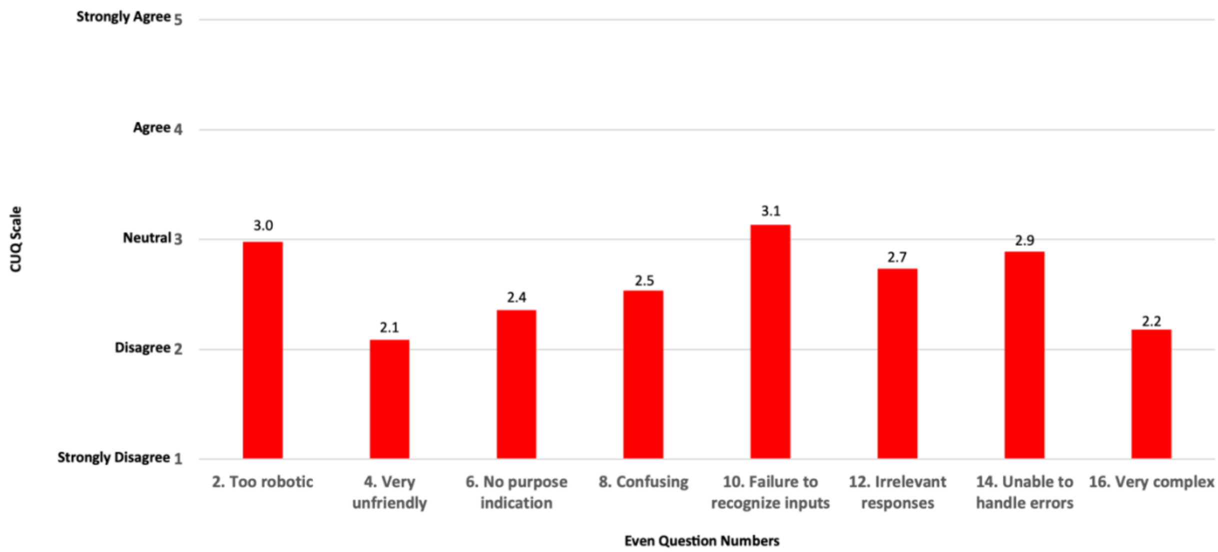
The mean chatbot usability score, measured by the CUQ, was 50.8 ± 4.0, and the median was 50. The odd-numbered questions of the CUQ represent positive aspects of the chatbot (**Figure 4.6**) ranked on a scale from 1–Strongly Disagree to 5–Strongly Agree. On this scale, Question 15, which states ‘The chatbot was very easy to use’ had the highest average ranking of 3.8, corresponding to Agree. The lowest average ranking was 3.1, corresponding to Neutral, for Questions 9 and 11, which state, ‘The chatbot understood me well’ and ‘Chatbot responses were useful, appropriate and informative.’

Figure 4.6. Average ranking for the positive aspects of Liv’s usability.



In **Figure 4.7**, the even-numbered questions of the CUQ represent negative aspects of the chatbot ranked on a scale from 1–Strongly Disagree to 5–Strongly Agree. Question 10, which states, ‘The chatbot failed to recognize a lot of my inputs’ had the highest average ranking of 3.1, corresponding to Neutral. The lowest average ranking was 2.1, corresponding to Disagree, for Question 4, which states, ‘The chatbot seemed very unfriendly.’

Figure 4.7. Average ranking for the negative aspects of Liv’s usability.



Qualitative Findings

Phase 4 interviews led to an understanding of patient experiences with Liv, including details about her personality, characteristics, escalation to human support, Liv's strengths and limitations, and noticing her improvement over time. The identified themes, descriptions, and sample quotes are summarized below (**Table 4.10**).

Table 4.10. Phase 4 interview thematic findings and sample interview quotes.

Theme	**Description of theme**	**Sample interview quotes**
The humanness of the chatbot	Participants commented on Liv's personality and characteristics.	"I loved the idea that she has a name, I love the name Liv, that felt very personable." (53)
		"It's cool to have a chat, when you don't have nobody else to talk to." (81)
		Participants described Liv as "pleasant" (25) and "upbeat" (25), although some did not appreciate this personality stating they

		"would prefer the neutral interaction" and "don't want AI to be [their] best friend." (11)
Convenience: knowing that chatbot would escalate	Participants emphasized the convenience of knowing that Liv would escalate unresolved questions to human support to receive real-time assistance.	"I liked being able to ask a question and her getting back to you saying, you know, that's a good question, I will get somebody to get back to you and find out about it. I liked being able to have her there and ask questions if you needed." (25)
		"It wasn't frustrating because I knew I'd be able to ask a real person the question, and deal with it that way." (23)
Mixed Experiences and Improvement Over Time	Participants shared both positive and negative experiences with Liv, noticing improvements in over time.	Some had positive experiences: "I used her a couple of times, and asked a few questions, and always got an answer. And she's easy to use and got back to me pretty quick. So I enjoyed it." (25)
		Others had negative experiences: "She didn't understand a word I said. I think she got better towards the end." (53) and "some of my questions were pretty specific and Liv didn't understand." (60)
		Many noticed her improvement: "I felt like she got better as the program went on. I just feel like things got better that she was more able to ascertain when I was really asking a question." (62) and "I think as she was getting refined, I think she was probably getting better." (34)

Question Content

After review of the 131 conversation logs that were escalated to agents for live support, six question content areas were identified. Many of these 131 conversation logs included multiple topics (n=34). Identified content areas, ordered from most frequently occurring to least frequently occurring, include information technology (IT), program content, program events, program feedback and social connection. A description of these content areas with frequencies is included in **Table 4.11**.

Table 4.11. Description of patient conversation topics through Liv.

Conversation topic	Frequency (# of conversations)	Description of topic
Information technology (IT)	78	Patient inquiries for website issues, glitches, individual device assistance, and website navigation questions.
Program content	44	Patient questions about energize or life skills program content (exercise, health education, psychology videos and activities) and program features (activity tracking, points, tracking log, community activity).
Program events	34	Patient inquiries about program events such as accessing group sessions and rescheduling one to one check-in calls.
General program feedback and research inquiries	12	Patient feedback to guide program iterations and specific inquiries about research administration (ex. surveys, program withdrawal, continued access, etc.).
Social connection	10	Patient inquires that included small talk and other social support with the live support team.

4.5. Discussion

Using a multi-phased, multi-method design, the study describes the development and evaluation of patient experiences with our FAQ chatbot ("Liv"). In our evaluation of acceptability, patients revealed an appreciation of the chatbot's personality, convenience knowing the chatbot could escalate to human agents, noticed Liv's potential for improvement over time, and highlighted Liv's current strengths/limitations. Usability evaluation revealed strengths in specific areas like Liv's ease of use, conversation navigation, and being welcoming at startup, with an overall below average CUQ usability score. Finally, user engagement analysis revealed 259 total conversations during deployment, a 33% resolution rate with patients escalating 131 conversations to live support, and highest engagement during the first 30 days of deployment.

To our knowledge, this is the first study that uses an approach to development that combines the Chatbot Development Life Cycle with a patient-oriented research framework to investigate the development of a FAQ chatbot to support patients participating in an online mental wellness program. While many digital health programs exist to manage chronic disease and mental health concerns, researchers have not investigated the inclusion of a chatbot in these programs to mitigate ongoing challenges such as limited user adoption, low engagement, and high attrition rates [41-44]. Moreover, few chatbots in this area have reported their approaches to development, with fewer having reported including patients in the development of their chatbots [76].

4.5.1. Experience with involving patients in the non-technical development of an FAQ chatbot through a multi-phased, multi-method, framework-based approach.

Throughout our multi-phased development process, we engaged patients as consultants [195] to direct training of the chatbot's knowledge bank, improve the ease of accessing the chatbot, and promote the chatbot's use. While our chatbot requires further refinement by prioritizing the needs and perspectives of our participants, our results support the creation of a chatbot that was acceptable, usable, and engaging in its first iteration. Consistent with our results, a study by Figueroa et al. on the user-centered design of a physical activity chatbot recognized that patient engagement can result in a better understanding of patient preferences, improve chatbot design, and increase the likelihood that patients will use and benefit from the chatbot [157]. Their study of 18 patients used a four-phase approach that included online patient interviews and patient chatbot prototype testing similar to our pre-deployment phases [157].

Moreover, our multi-phased approach was theoretically informed with guidance from established frameworks. The Chatbot Development Life Cycle is a useful framework that allowed us to map out the phases of our development approach before, during and after chatbot deployment and evaluation [70]. These steps were synergized with the SPOR patient engagement framework to allow us to prioritize and include patients from the outset throughout the development process [127]. The literature reporting theoretical guidance for health chatbot development is infrequent. While the chatbot development life cycle is often used to guide development in Computing Science settings, this has not been reported in health chatbot development to our knowledge. Maenhout et al. have used the person-based approach (PBA) as a theoretical framework to guide development [160]. This approach guides a stepwise process of intervention planning, optimization, and a multi-method process evaluation to ensure end-user needs and perspectives are embedded in the chatbot's development [160]. The PBA aims to build iterative, in-depth qualitative research into the development process [160]. Like experiences with Liv's development that also used a qualitative approach for development and a multi-method evaluation, these researchers reported that their framework-based approach allowed for creating a multi-phased process that led to insights on patient preferences that may have otherwise not emerged [160]. Attention to patient preferences and needs is known to increase the usability, uptake, and effectiveness of digital health technology [53, 56, 125, 126]. Thus, using a framework-guided approach to development in this study may have assisted the process of engaging patients in Liv's development and can ultimately result in creating more acceptable, usable, and engaging chatbots.

4.5.2. Chatbot Acceptability: Importance of Escalation and Personality

Initially, participants reported being largely unfamiliar with the roles of chatbots. This was unsurprising given our older adult participant demographic. Mesbah et al. examined seniors' acceptance of health chatbots. They reported that seniors have limited usage of current technologies, such as chatbots, relative to younger adults, due to not growing up using them [196]. Researchers suggest that this lack of experience can lead to greater skepticism and fear of using chatbot technology [196]. Despite these barriers, including hesitancy and unfamiliarity, patients reported that Liv was an acceptable means of receiving program support because of her ability to escalate conversations and her personality.

Beginning in Phase 1, participants emphasized the need to connect with human support if the chatbot failed. After deployment in Phase 4, Liv could only resolve 33% of patient inquiries due to reasons including trouble matching user intents to the correct question in the knowledge bank, patient typing/messaging habits (i.e. text message splitting into individual messages), and lack of chatbot training for specific patient questions (i.e. individual device assistance). In cases where questions were left unresolved, patients appreciated Liv's ability to connect users with human agents. Van der Goot and Pilgrim identified this same need for human support in their study of motivation to use chatbots, reporting that older adults valued human contact in this context and used chatbot communication as a "stepping stone for human contact" (connecting to a live agent) [197]. Other researchers, such as Følstad et al., studying chatbots in customer service environments have suggested that a chatbot's failure to resolve an inquiry may not be detrimental to user experience if the user is promptly provided with an opportunity to escalate to a human service representative [198]. Guided by our patient's advice, in this early iteration of Liv's development, the decision to allow for prompt escalation for unresolved inquiries was appreciated by patients in Phase 4.

In addition to escalation, participants appreciated Liv's personality and humanness, reporting positive experiences with congruence in both interview findings and CUQ responses. In interviews, patients described Liv as upbeat, pleasant, and human-like and enjoyed that she had a name. This paralleled results on the CUQ where patients reported she was 'welcoming at setup' and 'friendly.' This was similar to findings from Maenhout et al., where participants also valued the chatbot's human-like, friendly personality and appreciated that it had a name [160]. Other researchers have also recognized the importance of health chatbot personality development, prioritizing making their chatbot human-like [148] and empathetic [158]. Smestad et al. suggested that chatbot personality significantly positively affects user experiences with chatbots and that this personality should match the user group [199]. Figueroa et al. also reported that chatbot dialogue structure and personality are elements of chatbot design that must be addressed to avoid user frustration and reduced engagement [157]. Moreover, systematic review evidence from Kocaballi et al. states that chatbots with personality are reported to improve user satisfaction, user engagement and dialogue quality [200]. Like others, we also suggest that researchers should prioritize the development of personality and dialogue, alongside training the chatbot's knowledge bank, to create a more engaging experience for patients.

4.5.3. Chatbot Usability

Liv's first iteration had successful usability in specific areas such as conversation navigation and ease of use, but had limitations in understanding, input recognition, and error coping. Compared with the average benchmark score of 68 outlined by Holmes et al. [188], at 50.8, Liv's overall usability was below average. This may not be unexpected for a first version of a research-grade chatbot. There were, however, positive elements that also came out in both interviews and CUQ responses. Participants reported the ease of using Liv for program support. These findings align with literature studying older adults' use of chatbots that reports they find chatbots offer a smooth, easy-to-use interaction [197]. Other researchers, such as Brandzaeg et al., have also reported chatbot user feedback stating that chatbots are easy to use because they allow for direct support through the app and eliminate the need to navigate elsewhere on a device [61].

Participants reported positive and negative experiences with Liv's understanding, input recognition, and error coping in interviews and the CUQ. Participant experiences with this early iteration of Liv may have varied case-by-case depending on the question topic and specificity. Davis et al. encountered similar limitations and suggested that ensuring an adequate chatbot knowledge bank would improve a chatbot's ability to recognize questions and provide suitable answers [156]. This suggestion may be true as Phase 4 interview participants reported noticeable improvements in Liv's ability during the program as her knowledge bank was iterated. Liv's knowledge bank and natural language understanding have been identified as areas to improve Liv's overall usability in the future.

4.5.4. Chatbot User Engagement

Patients actively engaged with Liv throughout the deployment period to receive program support. The engagement was highest at the beginning of the deployment, with most users messaging Liv and connecting with live agents. This was expected as previous programs had the most inquiries at the beginning of the program while the online platform was unfamiliar to users. As the first evaluation of a chatbot in our application, our engagement measures indicated that the chatbot was a frequently used source for program support even when alternatives such as emails and check-in phone calls were available. Making comparisons of our chatbot's engagement to what exists in the literature is challenging for three reasons: 1) FAQ chatbots are most often implemented in education and commercial settings without evaluation studies on engagement, 2) studies that report chatbot engagement are chatbot interventions (such as a mental health support chatbot, or a physical activity coach), and 3) studies that evaluate chatbot intervention engagement are heterogeneous and lack a standard for reporting engagement.

4.5.5. Strengths

This study on Liv's development and evaluation had many strengths. Development followed a framework-based approach with guidance from the CIHR SPOR patient engagement framework [127] and the Chatbot Development Life Cycle [70]. Using this framework-based approach allowed us to engage patients in multiple phases and ensured their inputs were included in the development. Our multi-method approach led to qualitative and quantitative findings that fostered a deep understanding of Liv's first iteration that will be helpful in further iterations. Liv's knowledge bank, dialogue, and personality were also developed in consultation with our multidisciplinary research team. This ensured that chatbot content was evidence-based, the dialogue was user-friendly, and the personality was engaging for patients. Finally, this study is an example of formative work with thorough documentation of chatbot development. Our team conducted a scoping review before developing this chatbot that highlighted the lack of reporting of formative research in chatbot development [76]. Without thorough documentation of the development process, future researchers may be unable to replicate similar approaches for their research.

4.5.6. Limitations

First, while patient input was incorporated throughout development, there is still an opportunity to engage patients more actively in this research. Patient engagement organizations, like the Strategy for Patient-Oriented Research (SPOR) [121] and the Patient-Centered Outcomes Research Institute (PCORI) [201], view patient and public involvement in health research as existing on a spectrum. This spectrum has one end with minimal opportunities for patient input and on the other maximum collaboration and shared decision-making between researchers and patient partners [195]. Although this study did not formally evaluate the level of patient engagement in development, the extent of patient engagement in our project was likely at the consultation (1st) and involvement (2nd) level of the spectrum (**Figure 4.8**). Engaging patients in research can be more challenging and resource-intensive [141], but it was still prioritized despite this being a time-bound thesis project.

Figure 4.8. The spectrum of patient engagement [195].

In addition to the relatively low level of patient engagement, this study had several other limitations. First, Zendesk had limited data records to measure engagement beyond what was reported in our results section. While Zendesk made escalation to live support unproblematic for patients, Zendesk currently does not provide researchers access to chatbot conversation history unless escalated to live support. Furthermore, other chatbots with more advanced natural language understanding (NLU) may be available on the market than Zendesk's chatbot. For example, a recent evaluation by Abdellatif et al. of the NLU capabilities of various chatbots reported that IBM Watson is the best performing in terms of its intent classification and chatbot confidence scores [202].

Second, although a validated measure (i.e., CUQ) was used to evaluate chatbot usability [188], there was no validated measure for chatbot acceptability and engagement. Sample sizes were relatively small across all phases, and this made analysis of quantitative metrics challenging and limited only to descriptive statistics. Moreover, participants completed the CUQ four weeks into the online mental wellness program. This timeline was chosen because most questions and interactions with Liv would occur while users were unfamiliar with the program early on. However, this meant that CUQ data did not capture potential improvements to the chatbot as it was refined throughout the program. Therefore, we conducted end-of-program patient interviews to further our understanding of participant experiences with usability, acceptability, and engagement. End-of-program interviews balanced the single timepoint measurement of the CUQ and provided a deeper description of chatbot experiences and guidance for future development.

Third, as participants were already enrolled in the EMPOWER program, an online mental wellness program, they may have naturally been more technically proficient than the general population of older adults, limiting the generalizability of these findings. Moreover, some participants had past experiences with the PPP trial, a similar online mental wellness program. Therefore, they may have had a greater technological proficiency than that of the general population of older adults. This limitation is evident in participant demographics for phase 4 chatbot interviews which included participants with technical proficiencies ranging from 'moderate' to 'expert' with no participants reporting that they were 'sort of' or a 'little' proficient. This potentially limits the applicability of these findings to other projects with similar populations.

Finally, the overall usability, design, layout, and reliability of the EMPOWER website were improved over previous program iterations. This improvement may have naturally resulted in fewer issues and fewer patient questions. While this was beneficial from a user experience and intervention delivery standpoint, this may have reduced the need for a chatbot on the platform and, ultimately, reduced user engagement with the chatbot.

4.6. Conclusions

This study demonstrates that an FAQ chatbot is an acceptable and personable support for patients participating in an online mental wellness intervention. Recognizing that early in development, this population of older-adult patients was initially skeptical and hesitant to even use a chatbot for program support, these results indicate the possibility of implementing chatbots among older adults. While Liv is still early in her development and may not have been able to resolve as many patient queries as hoped, Liv was evidence of the active usage and uptake of a chatbot among this older-adult population. Patients were engaged with the chatbot, sending messages throughout the program and connecting with live support to receive answers when Liv failed. By involving patients in development and using a framework-guided approach, the acceptability, usability, and user engagement with chatbots may be improved. Future work should explore how

Liv can be refined from the conversation log and question content acquired during this evaluation. Further work is warranted to actively engage patients in developing and iterating chatbots. As chatbot development is a cyclic process and this is only the first iteration, more work is needed to understand Liv's capabilities and potential in this specific application.

CHAPTER 5: DISCUSSION

This thesis discusses the development and evaluation of a FAQ chatbot to support users of an online mental wellness program. Chapter 2 includes a literature review compiling information on chatbots, their applications in healthcare, chatbot development, and patient engagement to provide background for the remaining chapters. Chapter 3 includes a scoping review examining the development approaches of chatbots for lifestyle and mental wellness interventions and exploring the reported level of patient involvement in the development process. Finally, Chapter 4 discusses the developed frequently asked question (FAQ) chatbot, and evaluates acceptability, usability, and user engagement. To our knowledge, this thesis contains the first scoping review exploring the non-technical approaches for chatbot development in healthy lifestyle and mental wellness interventions. Additionally, to our knowledge, it includes the first study that uses an approach to development that combines the Chatbot Development Life Cycle with a patient engagement framework and uses that to evaluate an FAQ chatbot to support older-adult users through an online mental wellness program. Key findings from this work include (1) a summary of approaches to non-technical chatbot development and (2) usability, acceptability, and user engagement outcomes of our chatbot's first iteration.

5.1 Non-technical Approaches to chatbot development for an online mental wellness program

In Chapter 3, we presented a scoping review that explored the non-technical approaches to chatbot development, including the level of patient involvement in the process. We discovered a lack of comprehensive reporting of the development process, including limited detail on the nature of patient involvement. These findings were significant because researchers like Busse et al. have suggested that thoroughly reported work in health technology's development and assessment is important to guide future development [203]. Other researchers also highlight the importance of patient involvement in developing digital health interventions, including chatbots, which have been proposed as a potential solution to address challenges such as limited uptake, low engagement, and high rates of attrition that can hinder intervention effectiveness [41-44].

The insights gained in Chapter 3 were crucial in guiding the development of our chatbot "Liv" (Chapter 4). While we recognized the potential advantage of incorporating a chatbot in our online mental wellness program, we faced uncertainty regarding developing a user-friendly chatbot responsive to our users. Through our scoping review, we identified various non-technical approaches to chatbot development, including collaboration with patient partners, co-design workshops, patient interviews, prototype testing, Wizard of Oz (WoZ) procedure, and using literature review to gain evidence-based knowledge for development. These previously evaluated approaches offered us a range of options to plan the design of our chatbot across multiple phases.

Our review highlighted the disparity between expert recommendations that supported patient engagement and the limited reporting of patient engagement in chatbot development [76]. This disparity emphasized the need to prioritize patient involvement in developing our chatbot, and report how this process occurred. As a result of time pressures to launch the online mental wellness program and complete my master's, we could not carry out more intensive collaboration with patient partners via co-design workshops. That would have been the most involved way to include patients in development that was identified in the scoping review [76].

Instead, we engaged patients throughout our phasic approach via interviews to explore needs and preferences as well as prototype testing and usability walkthroughs.

5.2. Findings from the first iteration of the FAQ chatbot, Liv

The study's FAQ chatbot, Liv, underwent a multi-method evaluation to assess her acceptability, usability, and user engagement. This multi-method approach allowed us to understand patient priorities, Liv's strengths, limitations, areas for improvement, and user experiences. Additionally, this first iteration resulted in knowledge of frequently asked question topics from the conversation logs acquired during deployment.

At the outset of the project, we hypothesized that engaging patients in developing the FAQ chatbot would result in greater acceptability, usability, and user engagement. Recent literature, including a study by McCurdie et al., proposes that prioritizing patient engagement throughout development leads to a better understanding of patient needs, higher engagement levels, and improved intervention effectiveness [56]. Understanding that patient engagement offers numerous benefits, we sought to engage patients across multiple phases of our development process. This engagement commenced in Phase 1 when we recruited past participants of our mental wellness program to explore the need to include a chatbot in the program and establish patient priorities for the chatbot's development. In Phase 2, patients tested the prototype chatbot and suggested how to make refinements to increase ease of use (Phase 3A) and promote it on the program (Phase 3B). Finally, in Phase 4, we interviewed patients to understand their experiences with the chatbot.

Once the chatbot was evaluated in Phase 4, we found that even in its first iteration, it showed potential to be an acceptable, usable, and engaging means of support for patients participating in the online mental wellness program. These preliminary successes may in part be attributed to our involvement of patients in the development process. In a study by Mao et al. about user-centered design, they reported that engaging end-users in development through iterative prototyping and user testing cycles improved the ease of use and adoption of the technology [170]. In a similar study to ours that used a multi-phased, user-centered design approach to build a physical activity-promoting chatbot, researchers found that engaging patients in development resulted in a better understanding of their preferences, improved chatbot design, and increased the likelihood that patients would use and benefit from the chatbot [157]. In our case, areas that patients contributed to in development, such as improving the ease of use of the chatbot, were identified by patients as areas that were most successful during evaluation. Research by Gabrielli et al. in their study of the development of a chatbot to promote life skills among their adolescent patient population found that their co-design approach assisted with collecting patient needs and preferences, improving chatbot look and feel, and making improvements to unmet patient expectations in their prototype chatbot [150]. In our research, we also found that pre-deployment interviews and prototype testing also assisted with development of these same areas. This supports past literature, highlighting the importance of involving patients in developing digital health tools and chatbots [52-54, 150, 157, 170].

We also hypothesized that using frameworks to guide our chatbot's development played a role in its initial success. Other researchers, including van Gemert-Pijnen et al., have argued in their viewpoint paper that researchers should use theoretically-informed frameworks to improve the uptake and the impacts of digital health technology [165]. These researchers criticized current approaches to health technology development that lack a framework basis, reporting that these technologies fail to be meaningful, manageable, and sustainable for patients [165]. Moreover, researchers such as Maenhout et al. have found that using a framework-based, multi-method approach to develop their health promotion chatbot resulted in successful chatbot planning and optimization that was able to prioritize end user needs [160]. The guidance that frameworks provide for development [160, 165] can help mitigate challenges in involving patients in digital health development, such as 1) an underappreciation of the potential benefits of their involvement and 2) a limited understanding of how to get patients involved [141]. Our observations from Liv's development approach that combined elements from the steps from the Canadian Institutes of Health Research (CIHR) Strategy for Patient-Oriented Research (SPOR) framework [127] and the Chatbot Development Life Cycle [70] support the conclusions of previous research. The CIHR SPOR patient engagement framework provided meaningful ways to engage patients through its guiding principles (like inclusiveness and mutual respect) [127], which were were key in facilitating positive patient interactions during interviews and prototype testing.

Additionally, following the framework ensured that patient experiential knowledge was valued, and their inputs were translated into refining the chatbot's design. The Chatbot Development Life Cycle mapped out the chatbot development process. It allowed us to consider how we could include patients in the steps outlined in this cycle, namely in establishing chatbot requirements, architecture, deployment, monitoring, and promotion. The Chatbot Development Life Cycle guided us through a step-by-step process to form our multi-phased approach where patients were involved from the outset. Ultimately, without guidance from both frameworks, structuring the inclusion of patients into the chatbot's development would have been more challenging, and our preliminary successes may not have been achieved.

5.3. Limitations

This thesis has several limitations that need to be acknowledged. First, the scoping review outlined in Chapter 3 was constrained by the level of detail provided in the included studies. Due to the limited reporting, we could not definitively assess the extent of patient involvement in chatbot development. It is also important to recognize the inherent limitations of scoping reviews, including the potential for bias compared to systematic reviews and the absence of a quality assessment of included studies [172, 204]. However, given the heterogeneous nature of the current literature on chatbot development, a scoping review was the most appropriate method to map the existing research in this area.

Secondly, although we prioritized patient engagement in our chatbot's development process, the engagement level was limited to consulting and involving patients. Patient engagement organizations, like the SPOR [121] and the Patient-Centered Outcomes Research Institute (PCORI) [201], describe patient engagement as occurring on a continuum. Consultation includes asking patients for their opinions, advice and information to guide research; involvement means working directly with patients to understand their experiences and reflect those in research and development; and collaboration occurs when partnering with patients in every aspect of the research process [195]. Although engaging patients through collaboration is not always feasible due to its time and resource intensity [141], including patients as collaborators may have yielded important insights and increased engagement with our chatbot.

Lastly, it is important to recognize that this thesis is a time-bound project, only representing one iteration of chatbot development. Chatbot development is typically an ongoing, cyclical process that requires multiple iterations before achieving a fully autonomous and refined chatbot [70]. Though promising, the preliminary findings presented in this thesis should be viewed as the foundation for future development.

5.4. Future Directions & Conclusions

This thesis presents initial findings on the acceptability, usability, and user engagement of an FAQ chatbot designed to support patients in an online mental wellness program. While the chatbot succeeded in certain aspects, this thesis represents foundational work, with future efforts dedicated to improving overall usability and the user experience. The chatbot evaluation identified several areas for improvement, such as its intent recognition, response clarity, and error management. The data obtained from conversation logs in this initial phase will guide the chatbot's knowledge bank refinement to address these issues.

This thesis underscores the importance of patient involvement and the use of established frameworks in chatbot development. While previous research has touched upon development approaches, they often lack comprehensive reporting on patient involvement. To advance research in this area and create chatbots that meet the specific needs of patients, it is essential to conduct and document formative work guided by established frameworks, emphasizing patient engagement in the design process. This approach holds great promise in supporting participant adherence and ultimately improving outcomes for those with chronic disease.

REFERENCES

1. Adamopoulou, E. and L. Moussiades, *Chatbots: History, technology, and applications.* Machine Learning with Applications, 2020. **2**: p. 100006.
2. Zemčík, T., *A Brief History of Chatbots.* DEStech Transactions on Computer Science and Engineering, 2019.
3. Shum, H.-y., X.-d. He, and D. Li, *From Eliza to XiaoIce: challenges and opportunities with social chatbots.* Frontiers of Information Technology & Electronic Engineering, 2018. **19**(1): p. 10-26.
4. Hoy, M.B., *Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants.* Med Ref Serv Q, 2018. **37**(1): p. 81-88.
5. *What is Digital Health?* March 1, 2023]; September 22, 2020:[Available from: https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health.
6. Bates, M., *Health Care Chatbots Are Here to Help.* IEEE Pulse, 2019. **10**(3): p. 12-14.
7. Stoilova, E., *AI chatbots as a customer service and support tool.* ROBONOMICS: The Journal of the Automated Economy, 2021. **2**: p. 21-21.
8. Minutolo, A., et al., *A conversational agent for querying Italian Patient Information Leaflets and improving health literacy.* Computers in Biology and Medicine, 2022. **141**: p. 105004.
9. Yadav, D., et al., *Feedpal: Understanding opportunities for chatbots in breastfeeding education of women in india.* Proceedings of the ACM on Human-Computer Interaction, 2019. **3**(CSCW): p. 1-30.
10. Kataoka, Y., et al., *Development and early feasibility of chatbots for educating patients with lung cancer and their caregivers in Japan: mixed methods study.* JMIR cancer, 2021. **7**(1): p. e26911.
11. May, R. and K. Denecke. *Extending patient education with CLAIRE: an interactive virtual reality and voice user interface application*. in *Addressing Global Challenges and Quality Education: 15th European Conference on Technology Enhanced Learning, EC-TEL 2020, Heidelberg, Germany, September 14–18, 2020, Proceedings 15*. 2020. Springer.
12. Hess, G.I., G. Fricker, and K. Denecke, *Improving and evaluating eMMA's communication skills: a chatbot for managing medication.* Stud Health Technol Inform, 2019. **259**: p. 101-104.
13. Griffin, A.C., et al., *Information needs and perceptions of chatbots for hypertension medication self-management: a mixed methods study.* JAMIA open, 2021. **4**(2): p. ooab021.
14. Roca, S., et al., *A Study on the Impacts of Slot Types and Training Data on Joint Natural Language Understanding in a Spanish Medication Management Assistant Scenario.* Sensors, 2022. **22**(6): p. 2364.
15. Ghosh, S., S. Bhatia, and A. Bhatia, *Quro: facilitating user symptom check using a personalised chatbot-oriented dialogue system.* Stud Health Technol Inform, 2018. **252**: p. 51-56.
16. Espinoza, J., K. Crown, and O. Kulkarni, *A guide to chatbots for COVID-19 screening at pediatric health care facilities.* JMIR public health and surveillance, 2020. **6**(2): p. e18808.
17. Heald, B., et al., *Using chatbots to screen for heritable cancer syndromes in patients undergoing routine colonoscopy.* Journal of Medical Genetics, 2021. **58**(12): p. 807-814.
18. Echeazarra, L., J. Pereira, and R. Saracho, *TensioBot: a chatbot assistant for self-managed in-house blood pressure checking.* Journal of Medical Systems, 2021. **45**(4): p. 54.
19. Oh, Y.J., et al., *A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss.* International Journal of Behavioral Nutrition and Physical Activity, 2021. **18**: p. 1-25.
20. Perski, O., et al., *Does the addition of a supportive chatbot promote user engagement with a smoking cessation app? An experimental study.* Digital health, 2019. **5**: p. 2055207619880676.
21. Luo, T.C., et al., *Promoting physical activity through conversational agents: mixed methods systematic review.* Journal of Medical Internet Research, 2021. **23**(9): p. e25486.
22. Larbi, D., E. Gabarron, and K. Denecke. *Social Media Chatbot for Increasing Physical Activity: Usability Study*. in *pHealth*. 2021.

23. Huang, C.-Y., et al. *A chatbot-supported smart wireless interactive healthcare system for weight control and health promotion*. in *2018 IEEE international conference on industrial engineering and engineering management (IEEM)*. 2018. IEEE.
24. Abd-Alrazaq, A.A., et al., *An overview of the features of chatbots in mental health: A scoping review.* International Journal of Medical Informatics, 2019. **132**: p. 103978.
25. Denecke, K., S. Vaaheesan, and A. Arulnathan, *A mental health chatbot for regulating emotions (SERMO)-concept and usability test.* IEEE Transactions on Emerging Topics in Computing, 2020. **9**(3): p. 1170-1182.
26. Lee, Y.-C., N. Yamashita, and Y. Huang, *Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional.* Proceedings of the ACM on Human-Computer Interaction, 2020. **4**(CSCW1): p. 1-27.
27. Daley, K., et al., *Preliminary evaluation of the engagement and effectiveness of a mental health chatbot.* Frontiers in digital health, 2020. **2**: p. 576361.
28. Oh, K.-J., et al. *A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation*. in *2017 18th IEEE international conference on mobile data management (MDM)*. 2017. IEEE.
29. Sasseville, M., et al., *Digital health interventions for the management of mental health in people with chronic diseases: a rapid review.* BMJ Open, 2021. **11**(4): p. e044437.
30. Chatterjee, S., et al., *Diabetes structured self-management education programmes: a narrative review and current innovations.* The Lancet Diabetes & Endocrinology, 2018. **6**(2): p. 130-142.
31. *Prevalence of Chronic Diseases and Risk Factors among Canadians aged 65 years and older*. December 24, 2020 March 1, 2023]; Available from: https://www.canada.ca/en/services/health/publications/diseases-conditions/prevalence-chronic-disease-risk-factors-canadians-aged-65-years-older.html.
32. *About Chronic Diseases*. July 21, 2022 [cited 2023 February 27, 2023]; Available from: https://www.cdc.gov/chronicdisease/about/index.htm.
33. Roberts, K., et al., *Prevalence and patterns of chronic disease multimorbidity and associated determinants in Canada.* Health promotion and chronic disease prevention in Canada: research, policy and practice, 2015. **35**(6): p. 87.
34. *Noncommunicable diseases*. September 16, 2022 March 2, 2023]; Available from: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.
35. Angermayr, L., D. Melchart, and K. Linde, *Multifactorial Lifestyle Interventions in the Primary and Secondary Prevention of Cardiovascular Disease and Type 2 Diabetes Mellitus—A Systematic Review of Randomized Controlled Trials.* Annals of Behavioral Medicine, 2010. **40**(1): p. 49-64.
36. Oenema, A., et al., *Efficacy and Use of an Internet-delivered Computer-tailored Lifestyle Intervention, Targeting Saturated Fat Intake, Physical Activity and Smoking Cessation: A Randomized Controlled Trial.* Annals of Behavioral Medicine, 2008. **35**(2): p. 125-135.
37. Hoedjes, M., et al., *Effect of Postpartum Lifestyle Interventions on Weight Loss, Smoking Cessation, and Prevention of Smoking Relapse: A Systematic Review.* Obstetrical & Gynecological Survey, 2010. **65**(10): p. 631-652.
38. Gillies, C.L., et al., *Pharmacological and lifestyle interventions to prevent or delay type 2 diabetes in people with impaired glucose tolerance: systematic review and meta-analysis.* BMJ, 2007. **334**(7588): p. 299.
39. Ketola, E., R. Sipilä, and M. Mäkelä, *Effectiveness of individual lifestyle interventions in reducing cardiovascular disease and risk factors.* Annals of Medicine, 2000. **32**(4): p. 239-251.
40. Valenzuela, P.L., et al., *Lifestyle interventions for the prevention and treatment of hypertension.* Nature Reviews Cardiology, 2021. **18**(4): p. 251-275.
41. Whitelaw, S., et al., *Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review.* European Heart Journal - Digital Health, 2021. **2**(1): p. 62-74.

42. Yeager, C.M. and C.C. Benight, *If we build it, will they come? Issues of engagement with digital health interventions for trauma recovery.* Mhealth, 2018. **4**: p. 37.
43. Lie, S.S., et al., *Dropout from an eHealth intervention for adults with type 2 diabetes: a qualitative study.* Journal of medical Internet research, 2017. **19**(5): p. e187.
44. Price, M., et al., *Nonuse and dropout attrition for a web-based mental health intervention delivered in a post-disaster context.* Psychiatry, 2012. **75**(3): p. 267-284.
45. Boucher, E.M., et al., *Artificially intelligent chatbots in digital mental health interventions: a review.* Expert Review of Medical Devices, 2021. **18**(sup1): p. 37-49.
46. Dingler, T., et al., *The Use and Promise of Conversational Agents in Digital Health.* Yearb Med Inform, 2021. **30**(1): p. 191-199.
47. Chiru, A.L., I.A. Awada, and A.M. Florea. *A Support Process of Telemedicine Applications that Integrates a Chatbot*. in *2021 International Conference on e-Health and Bioengineering (EHB)*. 2021.
48. Kowatsch, T., et al., *Text-based healthcare chatbots supporting patient and health professional teams: preliminary results of a randomized controlled trial on childhood obesity.* 2017.
49. Gelbrich, K., J. Hagel, and C. Orsingher, *Emotional support from a digital assistant in technology-mediated services: Effects on customer satisfaction and behavioral persistence.* International Journal of Research in Marketing, 2021. **38**(1): p. 176-193.
50. Jain, M., et al. *Evaluating and informing the design of chatbots*. in *Proceedings of the 2018 designing interactive systems conference*. 2018.
51. Følstad, A., et al. *SIG: chatbots for social good*. in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018.
52. McCarthy, D., *Taking digital health to the next level: promoting technologies that empower consumers and drive health system transformation*. 2014: Commonwealth Fund.
53. Birnbaum, F., et al., *Patient engagement and the design of digital health.* Acad Emerg Med, 2015. **22**(6): p. 754-6.
54. Molina-Recio, G., et al., *Proposal for the user-centered design approach for health apps based on successful experiences: integrative review.* JMIR mHealth and uHealth, 2020. **8**(4): p. e14376.
55. Abras, C., D. Maloney-Krichmar, and J. Preece, *User-centered design.* Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, 2004. **37**(4): p. 445-456.
56. McCurdie, T., et al., *mHealth Consumer Apps: The Case for User-Centered Design.* Biomedical Instrumentation & Technology, 2012. **46**(s2): p. 49-56.
57. Watt, M., J.C. Spence, and P. Tandon, *Development of a Theoretically Informed Web-Based Mind-Body Wellness Intervention for Patients With Primary Biliary Cholangitis: Formative Study.* JMIR Formative Research, 2021. **5**(10): p. e29064.
58. Schlegell, K.V. and O. Abuomar. *Mental Health Frequently Asked Questions Chatbot Powered by Machine Learning*. in *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*. 2022.
59. Vaidyam, A.N., et al., *Chatbots and conversational agents in mental health: a review of the psychiatric landscape.* The Canadian Journal of Psychiatry, 2019. **64**(7): p. 456-464.
60. Dale, R., *The return of the chatbots.* Natural Language Engineering, 2016. **22**(5): p. 811-817.
61. Brandtzaeg, P.B. and A. Følstad. *Why People Use Chatbots*. 2017. Cham: Springer International Publishing.
62. Dokukina, I. and J. Gumanova, *The rise of chatbots–new personal assistants in foreign language learning.* Procedia Computer Science, 2020. **169**: p. 542-546.
63. Weizenbaum, J., *ELIZA—a computer program for the study of natural language communication between man and machine.* Commun. ACM, 1966. **9**(1): p. 36–45.
64. Molnár, G. and Z. Szüts. *The Role of Chatbots in Formal Education*. in *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)*. 2018.
65. *SMS: the language of 6 billion people*. 2015, PORTIORESEARCH.

66. *Chatbot Platforms*. 2023 [cited 2023; Available from: https://www.trustradius.com/chatbot.
67. Cameron, G., et al. *Back to the future: lessons from knowledge engineering methodologies for chatbot design and development*. in *British HCI Conference 2018*. 2018. BCS Learning & Development Ltd.
68. *How To Develop a Chatbot From Scratch*. May 8, 2017 [cited 2023; Available from: https://chatbotsmagazine.com/how-to-develop-a-chatbot-from-scratch-62bed1adab8c.
69. Perez-Soler, S., et al., *Choosing a chatbot development tool.* IEEE Software, 2021. **38**(4): p. 94-103.
70. Sheth, B. *The Bot Lifecycle*. May 13, 2016 [cited 2023; Available from: https://chatbotsmagazine.com/the-bot-lifecycle-1ff357430db7.
71. Griffin, A.C., et al., *Conversational Agents for Chronic Disease Self-Management: A Systematic Review.* AMIA Annu Symp Proc, 2020. **2020**: p. 504-513.
72. Oh, Y.J., et al., *A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss.* International Journal of Behavioral Nutrition and Physical Activity, 2021. **18**(1): p. 160.
73. Abd-Alrazaq, A.A., et al., *Effectiveness and Safety of Using Chatbots to Improve Mental Health: Systematic Review and Meta-Analysis.* J Med Internet Res, 2020. **22**(7): p. e16021.
74. Abd-Alrazaq, A.A., et al., *Perceptions and opinions of patients about mental health chatbots: scoping review.* Journal of medical Internet research, 2021. **23**(1): p. e17828.
75. Laranjo, L., et al., *Conversational agents in healthcare: a systematic review.* J Am Med Inform Assoc, 2018. **25**(9): p. 1248-1258.
76. Sadasivan, C., et al., *Examining Patient Engagement in Chatbot Development Approaches for Healthy Lifestyle and Mental Wellness Interventions: Scoping Review.* J Particip Med, 2023. **15**: p. e45772.
77. Fadhil, A., *Can a chatbot determine my diet?: Addressing challenges of chatbot application for meal recommendation.* arXiv preprint arXiv:1802.09100, 2018.
78. *The Top 12 Health Chatbots*. 2021 August 13, 2021 March 30, 2023]; Available from: https://medicalfuturist.com/top-12-health-chatbots/.
79. Santosh, K., et al., *Privacy, security, and ethical issues.* Artificial Intelligence and Machine Learning in Public Healthcare: Opportunities and Societal Impact, 2021: p. 65-74.
80. Gerke, S., T. Minssen, and G. Cohen, *Ethical and legal challenges of artificial intelligence-driven healthcare.* Artificial Intelligence in Healthcare, 2020: p. 295-336.
81. Fine, L.J., et al., *Prevalence of multiple chronic disease risk factors: 2001 National Health Interview Survey.* American Journal of Preventive Medicine, 2004. **27**(2, Supplement): p. 18-24.
82. *Chronic Disease Risk Factors*. 2023 January 25, 2023 April 1, 2023]; Available from: https://www.canada.ca/en/public-health/services/chronic-diseases/chronic-disease-risk-factors.html.
83. Bernell, S. and S.W. Howard, *Use Your Words Carefully: What Is a Chronic Disease?* Front Public Health, 2016. **4**: p. 159.
84. *Prevalence of Chronic Diseases Among Canadian Adults*. February 5, 2019 March 27, 2022]; Available from: https://www.canada.ca/en/public-health/services/chronic-diseases/prevalence-canadian-adults-infographic-2019.html.
85. Goodman, R.A., et al., *Defining and measuring chronic conditions: imperatives for research, policy, program, and practice.* Prev Chronic Dis, 2013. **10**: p. E66.
86. *Projected population, by projection scenario, age and sex, as of July 1 (x 1,000)*. 2022 August 22, 2022 [cited 2023 March 27, 2023]; Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710005701.
87. Tam, T., *Aging and chronic diseases: a profile of Canadian seniors.* 2021.
88. Canada, C.D.P.A.o. *2018 pre-budget submission to the House of Commons Standing Committee on Finance*. 2017 August 4, 2017 March 30, 2023]; Available from:

https://www.ourcommons.ca/Content/Committee/421/FINA/Brief/BR9073636/br-external/ChronicDiseasePreventionAllianceOfCanada-e.pdf.
89. Megari, K., *Quality of life in chronic disease patients.* Health psychology research, 2013. **1**(3).
90. Patrick, D.L. and P. Erickson, *Health status and health policy: quality of life in health care evaluation and resource allocation.* 1993.
91. Benkel, I., M. Arnby, and U. Molander, *Living with a chronic disease: A quantitative study of the views of patients with a chronic disease on the change in their life situation.* SAGE Open Med, 2020. **8**: p. 2050312120910350.
92. Hwu, Y.-J., *The Impact of Chronic Illness on Patients.* Rehabilitation Nursing, 1995. **20**(4): p. 221-225.
93. Elmslie, K., *Against the growing burden of disease.* Ottawa: Public Health Agency of Canada, Centre for Chronic Disease Prevention, 2012.
94. Braillard, O., et al., *The impact of chronic disease management on primary care doctors in Switzerland: a qualitative study.* BMC Fam Pract, 2018. **19**(1): p. 159.
95. Wilson, J.J. and K.M. Gil, *The efficacy of psychological and pharmacological interventions for the treatment of chronic disease-related and non-disease-related pain.* Clinical Psychology Review, 1996. **16**(6): p. 573-597.
96. Booker, R., *Chronic obstructive pulmonary disease: non-pharmacological approaches.* British Journal of Nursing, 2005. **14**(1): p. 14-18.
97. Bradley, T., et al., *Systematic review of lifestyle interventions to improve weight, physical activity and diet among people with a mental health condition.* Systematic Reviews, 2022. **11**(1): p. 198.
98. Ding, H., et al., *The effects of chronic disease management in primary health care: evidence from rural China.* Journal of Health Economics, 2021. **80**: p. 102539.
99. Foreyt, J.P., *Need for Lifestyle Intervention: How to Begin.* The American Journal of Cardiology, 2005. **96**(4, Supplement): p. 11-14.
100. Ricanati, E.H., et al., *Mitigating preventable chronic disease: progress report of the Cleveland Clinic's Lifestyle 180 program.* Nutrition & Metabolism, 2011. **8**: p. 1-12.
101. Manger, S., *Lifestyle interventions for mental health.* Australian Journal of General Practice, 2019. **48**(10): p. 670-673.
102. Ma, C., et al., *Effects of weight loss interventions for adults who are obese on mortality, cardiovascular disease, and cancer: systematic review and meta-analysis.* bmj, 2017. **359**.
103. Dunn, A.L., *Effectiveness of lifestyle physical activity interventions to reduce cardiovascular disease.* American journal of lifestyle medicine, 2009. **3**(1_suppl): p. 11S-18S.
104. Blokstra, A., I. van Dis, and W.M.M. Verschuren, *Efficacy of multifactorial lifestyle interventions in patients with established cardiovascular diseases and high risk groups.* European Journal of Cardiovascular Nursing, 2012. **11**(1): p. 97-104.
105. Ornish, D., et al., *Can lifestyle changes reverse coronary heart disease?: The Lifestyle Heart Trial.* The Lancet, 1990. **336**(8708): p. 129-133.
106. Kritchevsky, S.B., et al., *Intentional weight loss and all-cause mortality: a meta-analysis of randomized clinical trials.* PloS one, 2015. **10**(3): p. e0121993.
107. Willett, W., et al., *Chapter 44. Prevention of chronic disease by means of diet and lifestyle changes (Table 44.2).* Disease Control Priorities in Developing Countries. New York and Washington, DC: Oxford University Press and the World Bank, 2006.
108. Burgess, E., P. Hassmén, and K.L. Pumpa, *Determinants of adherence to lifestyle intervention in adults with obesity: a systematic review.* Clinical Obesity, 2017. **7**(3): p. 123-135.
109. Burgess, E., et al., *Behavioural treatment strategies improve adherence to lifestyle intervention programmes in adults with obesity: a systematic review and meta-analysis.* Clinical Obesity, 2017. **7**(2): p. 105-114.
110. Fappa, E., et al., *Lifestyle intervention in the management of metabolic syndrome: could we improve adherence issues?* Nutrition, 2008. **24**(3): p. 286-291.

111. Ronquillo, Y., A. Meyers, and S.J. Korvek, *Digital Health*, in *StatPearls*. 2023, StatPearls Publishing

Copyright © 2023, StatPearls Publishing LLC.: Treasure Island (FL).
112. Lowery, C., *What Is Digital Health and What Do I Need to Know About It?* Obstetrics and Gynecology Clinics of North America, 2020. **47**(2): p. 215-225.
113. Osipov, V.S. and T.V. Skryl, *Impact of digital technologies on the efficiency of healthcare delivery*, in *IoT in Healthcare and Ambient Assisted Living*. 2021, Springer. p. 243-261.
114. Bhavnani, S.P., *Digital Health: Opportunities and Challenges to Develop the Next-Generation Technology-Enabled Models of Cardiovascular Care.* Methodist Debakey Cardiovasc J, 2020. **16**(4): p. 296-303.
115. Lupton, D., *The digitally engaged patient: Self-monitoring and self-care in the digital health era.* Social Theory & Health, 2013. **11**: p. 256-270.
116. Gentili, A., et al., *The cost-effectiveness of digital health interventions: A systematic review of the literature.* Front Public Health, 2022. **10**: p. 787135.
117. Manteghinejad, A. and S.H. Javanmard, *Challenges and opportunities of digital health in a post-COVID19 world.* J Res Med Sci, 2021. **26**: p. 11.
118. Sharon, T., *The Googlization of health research: from disruptive innovation to disruptive ethics.* Personalized medicine, 2016. **13**(6): p. 563-574.
119. Vayena, E., et al., *Digital health: meeting the ethical and policy challenges.* Swiss medical weekly, 2018. **148**: p. w14571.
120. Solomon, D.H. and R.S. Rudin, *Digital health technologies: opportunities and challenges in rheumatology.* Nature Reviews Rheumatology, 2020. **16**(9): p. 525-535.
121. *Strategy for Patient-Oriented Research*. March 24, 2023 [cited 2023; Available from: https://cihr-irsc.gc.ca/e/41204.html.
122. Frisch, N., et al., *Patient-oriented research competencies in health (PORCH) for researchers, patients, healthcare providers, and decision-makers: results of a scoping review.* Research involvement and engagement, 2020. **6**(1): p. 1-14.
123. Herz, J., *Wearables are totally failing the people who need them most.* Wired, 2014.
124. Rosen, D., M. Stefanone, and D. Lackaff, *Proceedings of the 43rd Hawaii International Conference on System Sciences.* 2010.
125. Baker, T.B., D.H. Gustafson, and D. Shah, *How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research.* Journal of medical Internet research, 2014. **16**(2): p. e36.
126. Mao, J.-Y., et al., *The state of user-centered design practice.* Communications of the ACM, 2005. **48**(3): p. 105-109.
127. *Strategy for Patient-Oriented Research - Patient Engagement Framework*. Available from: https://cihr-irsc.gc.ca/e/48413.html.
128. Fergusson, D. and Z. Monfaredi, *Demonstrating the value of patient-oriented research in Ontario.* CMAJ, 2018. **190**(Suppl): p. S53-S54.
129. Staniszewska, S., et al., *The GRIPP checklist: Strengthening the quality of patient and public involvement reporting in research.* International Journal of Technology Assessment in Health Care, 2011. **27**(4): p. 391-399.
130. Shah, A., et al., *Digital Health Interventions for Depression and Anxiety Among People With Chronic Conditions: Scoping Review.* J Med Internet Res, 2022. **24**(9): p. e38030.
131. Escriva Boulley, G., et al., *Digital health interventions to help living with cancer: A systematic review of participants' engagement and psychosocial effects.* Psycho-Oncology, 2018. **27**(12): p. 2677-2686.
132. Widmer, R.J., et al., *Digital Health Interventions for the Prevention of Cardiovascular Disease: A Systematic Review and Meta-analysis.* Mayo Clinic Proceedings, 2015. **90**(4): p. 469-480.

133. Chatterjee, A., et al., *Digital Interventions on Healthy Lifestyle Management: Systematic Review.* J Med Internet Res, 2021. **23**(11): p. e26931.
134. Roberts, A.L., et al., *Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis.* Journal of Cancer Survivorship, 2017. **11**(6): p. 704-719.
135. Fadhil, A. and S. Gabrielli, *Addressing challenges in promoting healthy lifestyles: the al-chatbot approach*, in *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2017, Association for Computing Machinery: Barcelona, Spain. p. 261–265.
136. Ta, V., et al., *User Experiences of Social Support From Companion Chatbots in Everyday Contexts: Thematic Analysis.* J Med Internet Res, 2020. **22**(3): p. e16235.
137. Pereira, J. and Ó. Díaz, *Using Health Chatbots for Behavior Change: A Mapping Study.* Journal of Medical Systems, 2019. **43**(5): p. 135.
138. Oh, Y.J., et al., *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): p. 160.
139. Gray-Burrows, K.A., et al., *Role of patient and public involvement in implementation research: a consensus study.* BMJ Quality & Safety, 2018. **27**(10): p. 858.
140. O'Brien, N., et al., *Integrating Evidence From Systematic Reviews, Qualitative Research, and Expert Knowledge Using Co-Design Techniques to Develop a Web-Based Intervention for People in the Retirement Transition.* J Med Internet Res, 2016. **18**(8): p. e210.
141. Baines, R., et al., *Meaningful patient and public involvement in digital health innovation, implementation and evaluation: A systematic review.* Health Expectations, 2022. **25**(4): p. 1232-1245.
142. Arksey, H. and L. O'Malley, *Scoping studies: towards a methodological framework.* International Journal of Social Research Methodology, 2005. **8**(1): p. 19-32.
143. Levac, D., H. Colquhoun, and K.K. O'Brien, *Scoping studies: advancing the methodology.* Implementation Science, 2010. **5**(1): p. 69.
144. Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org.
145. *Age limits and adolescents.* Paediatr Child Health, 2003. **8**(9): p. 577-8.
146. McHugh, M.L., *Interrater reliability: the kappa statistic.* Biochem Med (Zagreb), 2012. **22**(3): p. 276-82.
147. Staniszewska, S., et al., *GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research.* BMJ, 2017. **358**: p. j3453.
148. Dhinagaran, D.A., et al., *Conversational Agent for Healthy Lifestyle Behavior Change: Web-Based Feasibility Study.* JMIR Form Res, 2021. **5**(12): p. e27956.
149. Gabrielli, S., et al., *Engagement and Effectiveness of a Healthy-Coping Intervention via Chatbot for University Students During the COVID-19 Pandemic: Mixed Methods Proof-of-Concept Study.* JMIR Mhealth Uhealth, 2021. **9**(5): p. e27965.
150. Gabrielli, S., et al., *A Chatbot-Based Coaching Intervention for Adolescents to Promote Life Skills: Pilot Study.* JMIR Hum Factors, 2020. **7**(1): p. e16762.
151. Greer, S., et al., *Use of the Chatbot "Vivibot" to Deliver Positive Psychology Skills and Promote Well-Being Among Young People After Cancer Treatment: Randomized Controlled Feasibility Trial.* JMIR Mhealth Uhealth, 2019. **7**(10): p. e15018.
152. Krishnakumar, A., et al., *Evaluating Glycemic Control in Patients of South Asian Origin With Type 2 Diabetes Using a Digital Therapeutic Platform: Analysis of Real-World Data.* J Med Internet Res, 2021. **23**(3): p. e17908.
153. Maher, C.A., et al., *A Physical Activity and Diet Program Delivered by Artificially Intelligent Virtual Health Coach: Proof-of-Concept Study.* JMIR Mhealth Uhealth, 2020. **8**(7): p. e17558.

154. To, Q.G., C. Green, and C. Vandelanotte, *Feasibility, Usability, and Effectiveness of a Machine Learning-Based Physical Activity Chatbot: Quasi-Experimental Study.* JMIR Mhealth Uhealth, 2021. **9**(11): p. e28577.
155. Piao, M., et al., *Use of the Healthy Lifestyle Coaching Chatbot App to Promote Stair-Climbing Habits Among Office Workers: Exploratory Randomized Controlled Trial.* JMIR Mhealth Uhealth, 2020. **8**(5): p. e15085.
156. Davis, C.R., et al., *A Process Evaluation Examining the Performance, Adherence, and Acceptability of a Physical Activity and Diet Artificial Intelligence Virtual Health Assistant.* Int J Environ Res Public Health, 2020. **17**(23).
157. Figueroa, C.A., et al., *Conversational Physical Activity Coaches for Spanish and English Speaking Women: A User Design Study.* Front Digit Health, 2021. **3**: p. 747153.
158. Issom, D.Z., et al., *Toward a Conversational Agent to Support the Self-Management of Adults and Young Adults With Sickle Cell Disease: Usability and Usefulness Study.* Front Digit Health, 2021. **3**: p. 600333.
159. Larbi, D., K. Denecke, and E. Gabarron, *Usability Testing of a Social Media Chatbot for Increasing Physical Activity Behavior.* J Pers Med, 2022. **12**(5).
160. Maenhout, L., et al., *Participatory Development and Pilot Testing of an Adolescent Health Promotion Chatbot.* Front Public Health, 2021. **9**: p. 724779.
161. Pecune, F., L. Callebert, and S. Marsella, *Designing Persuasive Food Conversational Recommender Systems With Nudging and Socially-Aware Conversational Strategies.* Front Robot AI, 2021. **8**: p. 733835.
162. Alghamdi, E. and R. Alnanih, *Chatbot Design for a Healthy Life to Celiac Patients: A Study According to a New Behavior Change Model.* International Journal of Advanced Computer Science and Applications, 2021. **12**(10).
163. Piao, M., et al., *Development and Usability Evaluation of a Healthy Lifestyle Coaching Chatbot Using a Habit Formation Model.* Healthc Inform Res, 2020. **26**(4): p. 255-264.
164. Wienert, J. and H. Zeeb, *Implementing Health Apps for Digital Public Health – An Implementation Science Approach Adopting the Consolidated Framework for Implementation Research.* Frontiers in Public Health, 2021. **9**.
165. van Gemert-Pijnen, J.E.W.C., et al., *A Holistic Framework to Improve the Uptake and Impact of eHealth Technologies.* J Med Internet Res, 2011. **13**(4): p. e111.
166. Islam, A. and B.M. Chaudhry, *A Relational Agent for the COVID-19 Patients: Design, Approach, and Implications.* JMIR Human Factors, 2022.
167. Kemp, E., et al., *Health literacy, digital health literacy and the implementation of digital health technologies in cancer care: the need for a strategic approach.* Health Promotion Journal of Australia, 2021. **32**(S1): p. 104-114.
168. Thabrew, H., et al., *Co-design of eHealth Interventions With Children and Young People.* Frontiers in Psychiatry, 2018. **9**.
169. Hanington, B. and B. Martin, *Universal methods of design expanded and revised: 125 Ways to research complex problems, develop innovative ideas, and design effective solutions*. 2019: Rockport publishers.
170. Mao, J.Y., et al., *The state of user-centered design practice - UCD is gaining industry acceptance - but its current practice needs fine-tuning.* COMMUNICATIONS OF THE ACM, 2005. **48**(3): p. 105-109.
171. *Workbook to guide the development of a Patient Engagement In Research (PEIR) Plan*. Available from: https://www.arthritisresearch.ca/wp-content/uploads/2018/06/PEIR-Plan-Guide.pdf.
172. Grant, M.J. and A. Booth, *A typology of reviews: an analysis of 14 review types and associated methodologies.* Health Information & Libraries Journal, 2009. **26**(2): p. 91-108.
173. Shaw, A., *Using Chatbots to Easily Create Interactive and Intelligent FAQ Webpages.* Journal of Applied Global Research, 2012. **5**(15).

174. Chatterjee, A., M. Gupta, and P. Agrawal. *FAQaugmenter: suggesting questions for enterprise FAQ pages*. in *Proceedings of the 13th international conference on web search and data mining*. 2020.
175. Han, S. and M.K. Lee, *FAQ chatbot and inclusive learning in massive open online courses.* Computers & Education, 2022. **179**: p. 104395.
176. Sethi, F., *FAQ (frequently asked questions) chatbot for conversation.* Authorea Prepr, 2020. **8**.
177. Rodsawang, C., et al., *Designing a competent chatbot to counter the COVID-19 pandemic and empower risk communication in an emergency response system.* OSIR Journal, 2020. **13**(2).
178. Walwema, J., *The WHO health alert: communicating a global pandemic with WhatsApp.* Journal of Business and Technical Communication, 2021. **35**(1): p. 35-40.
179. Ferreira, T.C., et al. *Ana: A brazilian chatbot assistant about covid-19*. in *Proceedings of Conference ACL 2020 Workshop on Natural Language Processing for COVID-19 (NLP-COVID)*. 2020.
180. Bharti, U., et al. *Medbot: Conversational artificial intelligence powered chatbot for delivering tele-health after covid-19*. in *2020 5th International Conference on Communication and Electronics Systems (ICCES)*. 2020. IEEE.
181. Peerani, F., et al., *A randomized controlled trial of a multicomponent online stress reduction intervention in inflammatory bowel disease.* Therapeutic Advances in Gastroenterology, 2022. **15**: p. 17562848221127238.
182. *Evidence Based Mental Wellness Programming Online for Adults Across Chronic Physical Conditions (EMPOwer)*. Available from: https://classic.clinicaltrials.gov/ct2/show/NCT05786482.
183. Dicicco-Bloom, B. and B.F. Crabtree, *The qualitative research interview.* Med Educ, 2006. **40**(4): p. 314-21.
184. Crowe, M., M. Inder, and R. Porter, *Conducting qualitative research in mental health: Thematic and content analyses.* Aust N Z J Psychiatry, 2015. **49**(7): p. 616-23.
185. Hsieh, H.F. and S.E. Shannon, *Three approaches to qualitative content analysis.* Qual Health Res, 2005. **15**(9): p. 1277-88.
186. Vaismoradi, M., H. Turunen, and T. Bondas, *Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study.* Nursing & Health Sciences, 2013. **15**(3): p. 398-405.
187. Saunders, B., et al., *Saturation in qualitative research: exploring its conceptualization and operationalization.* Qual Quant, 2018. **52**(4): p. 1893-1907.
188. Holmes, S., et al., *Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces?*, in *Proceedings of the 31st European Conference on Cognitive Ergonomics*. 2019, Association for Computing Machinery: BELFAST, United Kingdom. p. 207–214.
189. *Zendesk Explore: Reporting and Analytics*. [cited 2023; Available from: https://support.zendesk.com/hc/en-us/categories/4405298746906-Explore.
190. Tandon, P., et al., *Heal-me PiONEer (personalized online nutrition and exercise): An RCT assessing 2 levels of app-based programming in individuals with chronic disease.* Contemp Clin Trials, 2022. **118**: p. 106791.
191. Spanier, A. *About the bot builder*. 2023 [cited 2023; Available from: https://support.zendesk.com/hc/en-us/articles/4408838909210-About-the-bot-builder.
192. *How and When to Create an Escalation Path to a Human Chat Operator?* [cited 2023; Available from: https://conversation24.com/how-and-when-to-create-an-escalation-path-to-a-human-chat-operator/.
193. Zabój, D. *Fallback*. Apr 4, 2022 [cited 2023; Available from: https://www.chatbot.com/help/interactions/what-is-fallback-interaction/.
194. *Meet our virtual assistant, Liv*. Available from: https://vimeo.com/797483590/e72f07afa1?share=copy.

195. Shimmin, C. *PE-3: The Levels of Public and Patient Involvement*. 2016 Sep 21, 2016 [cited 2023; Available from: https://medium.com/knowledgenudge/pe-3-the-levels-of-patient-and-public-involvement-77026a547f2b.
196. Mesbah, N. and L. Pumplun, *"Hello, I'm here to help you"–Medical care where it is needed most: Seniors' acceptance of health chatbots.* 2020.
197. van der Goot, M.J. and T. Pilgrim. *Exploring age differences in motivations for and acceptance of chatbot communication in a customer service context*. in *International Workshop on Chatbot Research and Design*. 2019. Springer.
198. Følstad, A. and M. Skjuve, *Chatbots for customer service: user experience and motivation*, in *Proceedings of the 1st International Conference on Conversational User Interfaces*. 2019, Association for Computing Machinery: Dublin, Ireland. p. Article 1.
199. Smestad, T.L. and F. Volden. *Chatbot personalities matters: improving the user experience of chatbot interfaces*. in *Internet Science: INSCI 2018 International Workshops, St. Petersburg, Russia, October 24–26, 2018, Revised Selected Papers 5*. 2019. Springer.
200. Kocaballi, A.B., et al., *The Personalization of Conversational Agents in Health Care: Systematic Review.* J Med Internet Res, 2019. **21**(11): p. e15360.
201. Frank, L., et al., *The PCORI perspective on patient-centered outcomes research.* Jama, 2014. **312**(15): p. 1513-1514.
202. Abdellatif, A., et al., *A Comparison of Natural Language Understanding Platforms for Chatbots in Software Engineering.* IEEE Transactions on Software Engineering, 2022. **48**(8): p. 3087-3102.
203. Busse, R., et al., *BEST PRACTICE IN UNDERTAKING AND REPORTING HEALTH TECHNOLOGY ASSESSMENTS: Working Group 4 Report.* International Journal of Technology Assessment in Health Care, 2002. **18**(2): p. 361-422.
204. Pham, M.T., et al., *A scoping review of scoping reviews: advancing the approach and enhancing the consistency.* Research Synthesis Methods, 2014. **5**(4): p. 371-385.