CONVERSATIONAL AGENTS AND CONNECTED DEVICES TO SUPPORT CHRONIC DISEASE SELF-MANAGEMENT

Ashley C. Griffin

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Carolina Health Informatics Program in the Graduate School.

Chapel Hill
2021

Approved by:

Arlene E. Chung

Saif Khairat

Jaime Arguello

Yue Wang

Feng-Chang Lin

Stacy Cooper Bailey

Lucas Mentch

ABSTRACT

Ashley C. Griffin: Conversational Agents and Connected Devices
to Support Chronic Disease Self-Management
(Under the direction of Arlene E. Chung)

This dissertation focused on designing, developing, and evaluating the usability of a conversational agent for hypertension self-management. The objectives were to: 1) assess patient needs and preferences of a conversational agent; 2) design, develop, and evaluate a conversational agent prototype; and 3) identify physical activity clusters from wearable devices and evaluate the association between physical activity and health status, which could be used to facilitate future contextually aware dialogues as physical activity can be used to improve hypertension control.

Leveraging a user-centered design process, patients with hypertension (n=15) participated in semi-structured interviews to elicit needs and perceptions towards using conversational agents to assist with managing blood pressure and medications. Based on these needs, a functional prototype was iteratively designed and developed. Another sample of patients with hypertension (n=10) participated in task-based usability testing to assess the usability and acceptability for assisting with self-management tasks. Cluster analysis of wearable device data from patients (n=430) was conducted to identify physical activity patterns that could inform tailored coaching strategies. We examined the relationship between physical activity clusters and health status using cross-sectional and longitudinal analyses.

Usability testing revealed that patients demonstrated curiosity towards interacting with conversational agents for hypertension self-management behaviors for managing medications and refills, communicating with the care team, and maintaining healthy lifestyles. Patients expressed concerns about conversational agents being intrusive and providing too much information. Usability testing showed high rates of task completion and acceptability. Conversational user experience could be improved with additional navigational features of menu and back buttons, contextual error messages, and a health professional persona. Cluster analysis revealed three activity phenotypes of low, moderate, and high physical activity. Patients in the low activity cluster reported significantly worse patient-reported outcomes compared to those with moderate and high physical activity ($p<0.05$). The majority of patients remained in their original physical activity cluster across 6-month periods.

Within this emergent field, this research contributes towards improving the design, usability, and dialogues of self-management conversational agents. This research is an important step towards realizing the potential and implications of conversational agents to support chronic disease self-management and improve health outcomes.

To my parents, Scott and Clara Griffin. Thank you for the heartfelt support you have given me to make this dream possible.

ACKNOWLEDGEMENTS

I would like to give my sincerest appreciation to my advisor, Dr. Arlene Chung. Thank you for your endless determination and time spent helping me become a researcher. Your mentorship has been invaluable and shaped many aspects of my professional endeavors. I strive to have the amount of dedication and thoughtfulness as you do within the informatics community.

I am also incredibly grateful for the members of my dissertation committee, Drs. Saif Khairat, Stacy Bailey, Jaime Arguello, Yue Wang, Feng-Chang Lin, and Lucas Mentch. Thank you for your teaching, guidance, and thoughtful feedback. It has truly been a pleasure being able to learn from a multidisciplinary team of experts. I am very fortunate for the opportunity to work with each of you.

I am thankful for the support from the Carolina Health Informatics Program (CHIP) director, Dr. Javed Mostafa, who has provided incredibly meaningful direction and advice over the past four years. Thank you to all of the CHIP support staff, especially Lindsey Womack and Hannah David.

I would like to acknowledge the funding support from the National Library of Medicine Institutional Training Grant for Research Training in Biomedical Informatics and Data Science (T15-LM012500) which provided me with support for courses, workshops, conferences, publications, and research expenses. I would also like to acknowledge the North Carolina Translational and Clinical Sciences Institute pilot award as part of the National Institutes of Health's National Center for Advancing Translational Sciences (UL1TR002489) for funding a portion of this research (Aim 1 and 2). Support for research (Aim 3) was provided by the National Institutes of Health under award number R01EB025024 (PI Arlene Chung). This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. I am grateful for the professional and research services provided by the North Carolina Translational and Clinical Sciences Institute and the Odum Institute. I am also very appreciative of the participants who took part in these studies and were willing to share their opinions.

Thank you to my colleagues and friends for all of your support. I am thankful for the ability to meet so many creative and passionate colleagues and for the opportunity to embark in this journey together. I look forward to a lifetime of conversations, collaborations, and celebrations with you all.

Lastly, this achievement would not have been possible without the emotional encouragement from my family. To my parents and brother, thank you for your unconditional support and wisdom. Thank you to my husband for your endless positivity and advice throughout this endeavor.

TABLE OF CONTENTS

LIST OF TABLES xii

LIST OF FIGURES xiii

LIST OF ABBREVIATIONS AND SYMBOLS xiv

CHAPTER 1: INTRODUCTION 1

Chronic Disease Self-Management 1

Mobile Health Approaches for Chronic Disease Self-Management 5

Conversational Agents for Chronic Disease Self-Management 7

Challenges for Conversational Agents in Health Care 15

Research Aims 16

Organization of the Dissertation 19

CHAPTER 2: INFORMATION NEEDS AND PREFERENCES FOR A CONVERSATIONAL AGENT FOR HYPERTENSION SELF-MANAGEMENT 20

Background and Significance 20

Objectives 21

Methods 22

User-Centered Design Framework 22

Study Design 23

Study Setting and Participants 23

Recruitment 24

Interview Guide Development 24

Questionnaire Development 26

Interviews 27

Analysis 28

Results 29

Sample Characteristics 29

Information Needs for a Hypertension Medication Self-Management Chatbot 31

Perceptions and Perceived Use of a Chatbot 33

Barriers and Facilitators of Using a Chatbot 35

Discussion 38

Principal Findings 38

Implications for Health Care and Research 39

Limitations 40

Conclusion 41

CHAPTER 3: DESIGN, DEVELOPMENT, AND USABILITY OF A HYPERTENSION SELF-MANAGEMENT CHATBOT 43

Background and Significance 43

Objectives 45

Methods 45

Chatbot Design 45

Chatbot Development 46

Study Design 54

Sample and Sampling 55

Procedures 55

Analysis 57

Results 58

Participant Characteristics 58

Summary of Tasks 60
Discussion 65
Principal Findings 65
Implications for Health Care and Research 69
Limitations 70
Conclusion 71
CHAPTER 4: CLUSTERING PHYSICAL ACTIVITY PATTERNS FROM WEARABLE DEVICES 72
Background and Significance 72
Objectives 74
Methods 74
Study Setting 74
Data Preprocessing 75
Cluster Identification and Evaluation 77
Association between Physical Activity and Health Status 79
Results 80
Participant Characteristics 80
Clusters of Physical Activity Patterns 82
Association between Physical Activity and Patient-Reported Health Status 84
Discussion 86
Principal Findings 86
Implications for Health Care and Research 88
Limitations 89
Conclusion 90

CHAPTER 5: SUMMARY AND CONCLUSION 92
Summary of Findings 92
Implications for Health Care and Research 96
Directions for Future Research 100
Conversational Agents During COVID-19 and Beyond 103
Conclusion 105
APPENDIX A: SEMI-STRUCTURED INTERVIEW PROTOCOL 107
APPENDIX B: USABILITY TESTING PROTOCOL 119
REFERENCES 136

LIST OF TABLES

Table 1.1 Chronic disease self-management skills 2

Table 1.2 Text-based conversational agents for chronic disease self-management (2012-2018) 10

Table 2.1 Sample characteristics 29

Table 2.2 Self-management needs and desired features of a chatbot for hypertension self-management 31

Table 2.3 Perceptions and perceived frequency of use of a chatbot 34

Table 2.4 Barriers and facilitators of a chatbot for hypertension self-management 36

Table 3.1 Example training phrases and parameters by intent 52

Table 3.2 Sample characteristics 59

Table 3.3 Task summary 61

Table 3.4 Examples of participant utterances 62

Table 3.5 System Usability Scale scores 64

Table 3.6 Pervasive themes for usability strengths and shortcomings of Medicagent 65

Table 4.1 Sample characteristics 80

Table 4.2 Cluster profiles 83

Table 4.3 Patient-reported outcome scores across clusters 84

Table 4.4 Movement across clusters for consecutive disease activity scores 85

LIST OF FIGURES

Figure 3.1 System architecture 47

Figure 3.2 Screenshots of the user interface 50

Figure 3.3 Data extraction in Dialogflow 54

Figure 4.1 Sum of squared errors plot 79

Figure 4.2 Physical activity clusters 83

LIST OF ABBREVIATIONS AND SYMBOLS

AI	Artificial Intelligence
ANOVA	Analysis of variance
API	Application Programming Interface
app	Application
ASK-12	Adherence Starts with Knowledge 12
BMI	Body mass index
CCFA	Crohn’s and Colitis Foundation of America
CD	Crohn’s Disease
CDC	Centers for Disease Control and Prevention
COVID-19	Coronavirus disease-2019
ECA	Embodied conversational agent
EHR	Electronic Health Record
FHIR	Fast Healthcare Interoperability Resources
GAD	General Anxiety Disorder
HIPAA	Health Insurance Portability and Accountability Act
HIT	Health information technology
IBD	Inflammatory bowel disease
IC	Indeterminate colitis
IDEAS	Integrate, Design, Assess, and Share
iOS	iPhone operating system
ISO	International Organization for Standardization
JITAI	Just-in-time adaptive intervention

mHealth	Mobile health
ML	Machine learning
PGHD	Patient-generated health data
PHI	Protected health information
PHQ	Patient Health Questionnaire
PRO	Patient-reported outcome
PROMIS	Patient-Reported Outcomes Measurement Information System
RESET	Reveal, Escalate, Substitute, Explain, and Track
SCCAI	Simple Clinical Colitis Activity Index
SCDAI	Short Crohn’s Disease Activity Index
SD	Standard deviation
SE	Standard error
SMS	Short messaging service
SSE	Sum of squared errors
SUS	System Usability Scale
UC	Ulcerative Colitis
UTAUT	Unified Theory of Acceptance and Use of Technology
WHO	World Health Organization
a	Average distance from a data point to all data points in same cluster
b	Average distance from a data point to all data points in nearest cluster
c	Cluster centroid
C	Cluster

$dist$	Euclidean distance between two data points
i	Index of the cluster
K	Number of clusters
s	Silhouette coefficient
x	Data point belonging to a cluster
$\in$	Belongs to
$\sum$	Summation

CHAPTER 1: INTRODUCTION

Chronic Disease Self-Management

Sixty percent of U.S. adults suffer from chronic diseases which cost approximately 90% of the $3.5 trillion annual health care expenditures[1-3]. Chronic diseases are conditions that last over a year and require ongoing medical care[4]. The Centers for Disease Control and Prevention outlines four major risk factors for preventable chronic diseases: lack of physical activity, poor nutrition, excessive alcohol use, and tobacco use[4]. Hypertension is the most common chronic disease in the U.S., and the American College of Cardiology and American Heart Association recommend that blood pressure of at least 130/80 mmHg should be treated with lifestyle changes including physical activity and, in some patients, medication[5]. For physical activity, the American College of Cardiology and American Heart Association advises adults to obtain 90–150 minutes per week of aerobic activity[5]. Other types of exercise such as dynamic or isometric resistance training are also proven interventions for the prevention and treatment of hypertension[5]. Consuming a healthy diet, low sodium intake, adequate potassium, and low to moderate alcohol consumption are important dietary components for blood pressure control[5]. Appropriate nutrition and exercise can also facilitate weight loss which is a risk factor for a variety of chronic diseases[5]. Avoiding tobacco use and second-hand smoke also reduces risk for numerous chronic diseases. Because hypertension is often coexistent with other chronic diseases, such as diabetes mellitus and hypercholesterolemia, modifying the aforementioned lifestyle behaviors and adhering to pharmacologic interventions are important for managing a wide range of chronic diseases.

Lifestyle changes and ongoing self-management of chronic diseases are challenging and requires sufficient knowledge of the disease and the necessary skills to prevent or manage complications[6]. Self-management refers to an individual's ability to manage symptoms, treatments, physical and psychosocial consequences of the disease, and lifestyle changes[6]. Essential self-management skills focus on helping individuals gain confidence in controlling their symptoms and learning how health conditions impact their day-to-day life (Table 1.1)[7]. Learning these skills helps to continue normal routines, manage conditions, and handle any negative emotions that may arise[7, 8]. Although many of these self-management skills are shared across various diseases and populations, effective interventions require tailoring based on individual needs and values[8]. Successful chronic disease self-management is vital to achieve improved health outcomes, quality of life, and cost-effective care[9].

Table 1.1 Chronic disease self-management skills

Self-Management Skill	**Description**
Community resources	Finding and using community resources
Decision making	Making day-to-day choices about when to seek medical help or which treatments to try based on having enough and appropriate information
Maintaining a healthy lifestyle	Sustaining healthy habits such as physical activity, nutrition, sleep, and stress management
Managing medications	Using medications safely and effectively while minimizing side effects
Managing symptoms	Coping with symptoms such as pain or fatigue

Participating in social activities	Interacting with other people through social groups or events
Problem solving	Identifying and generating effective solutions in response to the disease becoming better or worse
Talking with friends and family	Communicating with friends, family members, and/or caregivers about the illness
Working with care team	Communicating with the care team about the illness

While self-management focuses on an individual’s ability to monitor their condition, it is influenced by complex, multifaceted components which are not solely at the individual level. The World Health Organization Adherence Framework (2003) outlines five interacting dimensions affecting adherence to treatments and lifestyle behaviors: socioeconomic, condition, therapy, health system, and patient-related factors[9]. Socioeconomic factors that impact adherence include age, race, education, income, employment status, and literacy[9]. Recent evidence suggests there are also technology-based disparities for health tracking and monitoring, which has been referred to as the “digital divide”[10, 11]. Condition-related factors impacting treatment adherence comprise comorbidities, level of disability (physical or psychosocial), and severity of symptoms[9]. Some chronic diseases are asymptomatic (e.g., hypertension), and patients may not perceive immediate benefits from lifestyle modifications or taking medications[12]. Therapy-related factors refer to the complexity of the treatment regimen, duration of treatment, and side effects[9]. Simplifying medication regimens, such as taking a medication once per day instead of multiple times per day, has been shown to improve adherence[13]. Health system factors involve the patient-physician relationship, length of patient consultations, community support resources, and reimbursement mechanisms[9]. Physician empathy has demonstrated improved patient trust, motivation, and treatment compliance[14]. Lastly, patient-related factors impacting treatment adherence are the resources, knowledge, beliefs, and treatment expectations of the patient[9]. For example, inadequate knowledge, lack of motivation, forgetfulness, or negative beliefs about the treatment efficacy often negatively impact adherence[9]. Among these five dimensions, patient-related factors are the most modifiable, and thus, are the focus of many chronic disease self-management interventions[9]. However, successfully executing lifestyle modifications and treatment adherence requires targeting multiple dimensions[9].

Traditional in-person chronic disease self-management programs have transitioned from strictly didactic interventions (1970s and 1980s) towards collaborative, theoretically grounded patient empowerment programs[15]. These multifaceted approaches move beyond patient education to provide feedback, reinforcement, and facilitation from health professionals or peer support groups[15-17]. Many types of interventions have found improvements in self-management behaviors through combinations of exercise programs, group discussions, individual plans (e.g., diet), instruction from health professionals (e.g., medication usage), or educational booklets[6, 15, 18, 19]. A wide range of self-management skills have been used in the aforementioned interventions to build confidence and self-efficacy, such as problem solving, decision making, goal setting, and resources. Overall, traditional chronic disease self-management programs have also found short-term improvements in depression, disability, pain, fatigue, health-related quality of life, and increased patient communication with care teams[18, 19]. Despite these positive health outcomes, face-to-face chronic disease self-management programs are resource-intensive, only available during scheduled sessions, and may be difficult to scale across larger populations.

Mobile Health Approaches for Chronic Disease Self-Management

Mobile health (mHealth) refers to medical and public health practice supported by mobile devices such as mobile phones, monitoring technologies, and other wireless devices[20]. With the rise in personal computing power and expansion of mobile connectivity, numerous mHealth technologies have emerged and have the potential to profoundly transform health care[21]. In the U.S., 96% of adults owned a mobile phone and 81% owned a smartphone in 2019[22]. Given their near ubiquity, portability, and low cost, mHealth applications and devices offers opportunities for patients to regularly access, track, and share health information with their care team. This ability to capture and transmit patient-generated health data (PGHD), which are health data that are created, recorded, or gathered from patients or their caregivers, can be incorporated into clinical care to facilitate patient-centered care and empowerment[23]. The use of PGHD has also demonstrated value in accelerating clinical, public health, and research insights[23-25]. As self-management occurs during day-to-day activities and generally outside of the clinical setting, mHealth devices have been used to supplement or be an alternative to traditional in-person self-management programs. The uses of mHealth for chronic disease self-management are vast and continuously evolving, and common mHealth approaches utilize smartphone applications (apps), patient portals, wearable devices, and short messaging service (SMS) text messaging.

Apps can support self-management through education, communication, and reminders, and thereby, improve outcomes including hemoglobin A1c, blood pressure, and medication adherence[26-30]. Patient portals, which are mobile or web-based platforms for patients to access their health information from a health care organization's electronic health record, have also been increasingly used to provide education, health tracking, and feedback from health professionals[31]. However, there is limited evidence of patient portals improving health outcomes, which may be due to the lack of patient-centered design and components that facilitate self-management such as goal setting, consistent monitoring, and timely feedback based on patient goals and status[31, 32].

Accelerometer-based wearable activity trackers (wearables) can promote healthy lifestyles with improvements in steps, moderate to vigorous physical activity, and weight loss[33, 34]. Less commonly available, wearable biometric sensors found in skin patches, earphones, and clothing may also promote healthy lifestyles and yield valuable medical-grade information[21]. Wireless or Bluetooth devices have been used to connect to mobile phones such as blood pressure monitors, weight scales, or electrocardiograms. Many of these wireless devices have been paired with lifestyle education and communication interventions for integrated health monitoring[27]. For example, mobile phone-based telemonitoring systems have been integrated within care settings where health professionals were alerted when data was outside of a specified range[35, 36]. As more data are generated in real-time through wireless or connected devices, recognition of patterns in health-related behaviors, such as physical activity, could allow interventions to provide valuable self-monitoring information and personalized feedback.

Thus far, short message service (SMS) text messaging interventions are the most common and successful mHealth approach for chronic disease self-management[37]. SMS interventions can be available on a basic cell phone or smartphone, used in areas without broadband or Wi-Fi access, and may be accessible for those with low health or technical literacy[37]. SMS approaches have shown improvements across a range of self-management behaviors, such as medication adherence, smoking cessation, and physical activity[38-42]. Many interventions have also been paired with wearables or Bluetooth-enabled devices in order to provide feedback on PGHD[37, 43]. Users have largely reported SMS interventions to be highly acceptable and easy to use, though many of these systems have generic messages and are unable to understand natural language inputs from users[26, 37]. This is specifically where more advanced conversational interfaces have potential advantages for tailored, automated two-way communication based on user inputs and preferences.

Conversational Agents for Chronic Disease Self-Management

The use of conversational agents for mHealth is an emerging area of research within the past decade[44]. Conversational agents are systems that can communicate with users in natural language through text or speech[45]. Conversational agents typically fall into two categories: 1) chatbot systems that mimic “chat” characteristics of human-human interaction with extended conversation abilities, and 2) task-oriented dialogue systems that complete specific tasks and are designed for shorter conversations[45]. Chatbots can mirror a therapeutic process, such as cognitive behavioral therapy or brief motivational interviewing[46]. The first well-established chatbot, ELIZA, simulated a Rogerian psychotherapist using pattern matching and keywords to respond to a user with open-ended messages[47]. Rogerian psychology later gave rise to cognitive behavioral therapy and brief motivational interviewing[48]. These processes promote goal setting, positive feedback, self-monitoring, overcoming obstacles, and education, which are key self-management components.

Task-oriented dialogue systems have become popular in smartphones, cars, and home controllers (i.e., Apple Siri, Amazon Alexa, Google Assistant). For example, these systems can control appliances, make calls or texts, or find nearby facilities. They offer convenience due to the ability to be deployed on familiar platforms, such as Facebook Messenger, Telegram, or Amazon Alexa, and can work across multiple devices and platforms. Within the past five years, software developer kits for conversational interfaces have been released to provide development tools and resources, such as Amazon Lex[49], Google Dialogflow[50], and Microsoft Bot Framework[51]. Given this new ability to develop and deploy conversational agents on existing platforms, we conducted a review of health-related voice apps for Amazon Alexa and Google Assistant to assess the current landscape of these commercial apps (2015–2017)[52]. We identified 309 voice apps that were targeted towards health or fitness, and the majority were available through Amazon Alexa. Apps were broadly focused on health education or fitness training, and few were targeted specifically towards patients, caregivers, or health professionals. This suggests there is a rapidly developing market of conversational agents for health, and it does not currently focus on chronic conditions or self-management.

Conversational agents have unique characteristics that make them highly suitable for delivering self-management skills. As behavior change is often facilitated by social support, conversational agents have the ability to provide empathic support and accountability[53]. Affective conversational agents, which show sympathy or active listening, can help users experiencing negative emotions that are often present with chronic diseases[54]. Several ELIZA-like chatbots that demonstrate empathy and mimic therapeutic processes have shown promise for improving health-related behaviors[55, 56]. These human sentiment-related interaction norms are present in human-agent interactions, and humans are twice as likely to respond adversely when faced with a negative utterance by the agent as compared to a human[57]. This can be explained through the Computers are Social Actors theorem which demonstrates how people have a natural propensity for interacting with computers as if they were people[58, 59]. People also perceive computers as more likable when flattered or humored by them[58, 59]. In particular, text-based conversational agents can engage users in more personal or stigmatized topics as compared to spoken queries[60]. Due to the sensitive nature of managing one's health (i.e., weight management, diet, physical activity, medications, tobacco and alcohol use), text-based conversational agents may provide the optimal medium to deploy self-management interventions.

Given the potential to deliver self-management components through a conversational agent, we conducted a systematic literature review to assess how text-based conversational agents have been used for chronic disease self-management[61]. Our review focused on all types of chronic diseases and older adult populations[62] as self-management is similar across conditions, such as managing medications effectively, and individuals often suffer from multiple chronic diseases. We found 12 studies that contained primary research findings for text-based conversational agents focused on chronic disease self-management (Table 1.2). The majority of the conversational agents were targeted towards mental health and focused on self-management through maintaining a healthy lifestyle (i.e., exercise, nutrition, sleep). Few provided community resources or assistance with managing medications. A small number of studies used established design principles, such as participatory or user-centered design, and none used heuristic evaluation.

Similar to other reviews of conversational agents for health[52-55], these 12 studies contained small sample sizes and short study durations. Study outcomes were largely focused on usability of conversational agents, and participants mostly reported positive attitudes with some concerns for privacy and shallow content. In the studies that examined patient-reported outcomes, there were statistically significant improvements on the Patient Health Questionnaire, Generalized Anxiety Disorder Scale, Perceived Stress Scale, Flourishing Scale, and Overall Anxiety Severity and Impairment Scale between the conversational agent intervention and control groups. This early evidence suggests text-based conversational agents are acceptable, usable, and may be effective in supporting self-management. However, the lack of methodological rigor and heterogeneity across study designs may limit generalizability. Future studies of conversational agents should prioritize established design principles and standardized evaluation metrics, such as the International Organization for Standardization (ISO) 9241-11 standards, to assess usability[63].

Table 1.2 Text-based conversational agents for chronic disease self-management (2012-2018)

Study	**Chronic Disease/ Population**	**Self-Management Skills**	**Outcomes**
Baskar et al, 2015[64]	Older Adults	• Maintaining a healthy lifestyle	• Attitudes, perceptions • Usage

Elmasri et al, 2016[65]	Substance Use	• Maintaining a healthy lifestyle	• Attitudes, perceptions • Client Satisfaction Survey • Usage
Fitzpatrick et al, 2017[66]	Depression	• Managing symptoms	• Patient Health Questionnaire (PHQ)-9 ($p=0.017$) • General Anxiety Disorder (GAD)-7 Scale ($p=0.004$) • Positive Affect Schedule ($p=0.951$) • Negative Affect Schedule ($p=0.80$) • Satisfaction ($p<0.001$) • Usage • Knowledge
Gaffney et al, 2013[67]	Depression	• Problem solving	• Resolution ratings ($p<0.05$) • Helpfulness ratings ($p<0.05$) • Depression and Anxiety Stress Scale ($p=0.36$) • Distress ($p=0.13$) • Usage
Kazemi et al, 2014[68, 69]	Substance Use	• Maintaining a healthy lifestyle • Community resources • Participating in social activities	• Attitudes, perceptions
Ly et al, 2017[70]	Depression	• Maintaining a healthy lifestyle	• Flourishing Scale ($p=0.032$) • Perceived Stress Scale ($p=0.048$) • Attitudes, perceptions • Usage
Schroeder et al, 2018[71]	Depression	• Managing symptoms • Problem solving	• Overall Anxiety Severity and Impairment Scale ($p<0.05$) • PHQ-9 ($p<0.01$) • System Usability Scale • Usage
Stein et al, 2017[72]	Diabetes	• Maintaining a healthy lifestyle	• Weight loss • Healthy meals logged • Satisfaction • Usage

Tsiourti et al, 2014[73]	Older Adults	• Maintaining a healthy lifestyle • Working with the care team • Talking with friends and family • Managing medications	• Attitudes, perceptions
van Heerden et al, 2017[74]	HIV/AIDS	• Maintaining a healthy lifestyle • Managing symptoms • Working with the care team • Talking with friends and family	• Attitudes, perceptions • Usage
Wang et al, 2018[75]	Substance Use	• Maintaining a healthy lifestyle • Talking with friends and family • Participating in social activities	• Smoking cessation • Usage
Watson et al, 2012[76]	Overweight/ obesity	• Maintaining a healthy lifestyle • Problem solving	• Step count (p=0.07) • Weight • Body Mass Index (p=0.44) • Self-efficacy • Attitudes, perceptions • Usage

Similar to other mHealth interventions, conversational agents vary considerably, and it is unclear how they differ based on patient characteristics and social determinants of health. Patients may need varying levels of support depending on sociodemographic characteristics, health status, and cultural factors, and these needs largely remain implicit and unaddressed in the existing body of research[77, 78]. For example, patients with low intrinsic motivation may benefit more from human support than those motivated to work on their own, and understanding user motivation could contribute to the optimal timing and type of support provided by conversational agents[78]. Future systems should consider the timing and nature of support as well as how individual characteristics affect one's ability to self-manage and use conversational technologies.

Only a few agents have been paired with wearables or apps to provide tailored information or praise based on changes in activity level, weight, or dietary habits[72, 76, 79]. Beyond conversational agents, previous studies have used wearables or sensors to personalize behavioral coaching strategies, which resulted in improvements in physical activity, dietary habits, weight loss, sleep, and clinical biomarkers (e.g., lipids, hemoglobin A1c levels)[80-84]. These behavioral just-in-time adaptive interventions (JITAI) are characterized by behavioral support corresponding to a need in real-time; content or timing of support is adapted according to data inputs; and support is triggered by the system and not directly by the user[82, 85]. For example, a JITAI could automatically detect changes in physical activity patterns using wearable data and deliver personalized support based on the user's GPS or weather data in real-time. These interventions often leverage digital phenotyping algorithms, which use data generated by smartphones and other connected devices to measure health or functioning[86]. Few conversational agents utilize JITAI approaches, though they offer a promising medium to deliver context-aware self-management dialogues based on the setting or functioning of the individual in real-time.

Thus far, the majority of health-related conversational agents have elementary dialogue management systems, which determine which action to take given the user's input and current state of the dialogue[61, 87]. Conversational agents are characterized by three main types of dialogue systems: 1) rule-based, 2) statistical data-driven, and 3) end-to-end neural[87]. Rule-based systems take the user through a sequence of predetermined steps where the system controls the dialogue by prompting and confirming user inputs (which are typically limited to a single word or phrase)[87]. Among existing conversational agents focused on health, rule-based systems are commonly used to administer survey instruments or assess symptoms[64, 65]. This is due, in part, to risks of the system not accurately understanding the user's input and providing inappropriate medical information[88]. Confirming or constraining user input to menu choices could mitigate some patient safety concerns, but there may be a tradeoff for improving the user experience and long-term engagement.

Designing rules for every type of possible interaction is challenging, and statistical data-driven dialogue systems move beyond hand-crafted rules to machine learning from data to address these challenges[87]. These types of systems leverage corpus-based or example-based data to train the natural language understanding engine, which uses probabilistic modeling to recognize the user's utterance and then route them to subsequent conversational nodes. Several health-related conversational agents have employed data-driven dialogue systems or used them in combination with rule-based approaches[70, 72, 89]. In contrast to retrieving responses from corpora or example datasets, end-to-end neural dialogue systems use deep learning to respond to answers that are not specifically in the training data. These neural models use a sequence-to-sequence approach which treat user inputs as a sequence and utilize deep learning to process and output a sequence one token at a time[87]. Neural models can generalize to new questions or phrases but require large amounts of training data and are primarily in the early stages of research and development for text-based dialogues. Therefore, neural systems are not often used for health use cases. Both statistical data-driven and end-to-end neural systems use conversational artificial intelligence (AI), which refers to techniques for creating software agents that can engage in natural conversational interactions[90]. The field of conversational AI currently faces challenges in the lack of available conversational datasets, natural language understanding for multiturn dialogues, evaluation of the quality of interactions, and personalization[87]. Thus, the field is largely still in its infancy, particularly for health use cases, and remains quite far from engaging with users in a truly natural conversational dialogues.

Challenges for Conversational Agents in Health Care

While conversational agents have become widely utilized across other industries, their use in health care presents a number of distinct privacy, regulatory, safety, and ethical challenges[91]. Many conversational developer platforms allow organizations to enter into a Health Insurance Portability and Accountability Act (HIPAA) Business Associate Agreement to support protection of health-related data[51, 92, 93]. Currently, Amazon Alexa is the only platform that provides a HIPAA eligible environment to build apps that transmit and receive protected health information though it is only available to select developers[94]. Several voice apps are currently operating under Amazon Alexa's HIPAA eligible environment. For example, these apps allow users to query their blood sugar measurements, locate an urgent care facility, schedule appointments, and check the status of home delivery prescriptions[94]. The Food and Drug Administration has not yet made an explicit statement regarding the level of enforcement for health-related conversational agents[95]. It is likely that conversational agents would be regulated similarly to mobile apps, which are not considered medical device manufacturers unless the app delivers care or makes care decisions.

In addition to regulatory and legal provisions, reimbursement mechanisms and the ability to incorporate actionable information from these agents into the clinical workflows remains underexplored. Safety protocols that establish monitoring, alerts, and escalation are needed prior to leveraging conversational agents in clinical care, particularly for those with unconstrained natural language input[91]. If data generated from conversational agents are integrated within Electronic Health Records (EHRs), summarizing data in a meaningful way is also essential to avoid burdening already time-constrained clinicians. Additional considerations for transparency and connection to health professionals when needed or requested are important for patient safety and trust. Ethical concerns include the use of assistive technologies, conversing with a system that mimics a human, or over-reliance, which is especially important for older adults who may need to maintain current levels of independence[73]. As with all technologies, careful consideration should be made to promote inclusive designs and prevent further exacerbating health disparities. Ensuring representative patient groups are involved in the design and algorithm training, including vernacular, tone, and contextual data is essential. Collectively, these complex factors must be thoughtfully studied to realize the potential of conversational agents to support self-management and improve health outcomes.

Research Aims

The purpose of this dissertation is three-fold: 1) to design and develop a conversational agent to support hypertension self-management and treatment adherence, 2) to evaluate the usability of this conversational agent, and 3) to identify physical activity patterns from wearable devices and evaluate their association with health status, which could be used to tailor the conversational agent. This research utilizes user-centered design methods, in which patients are involved throughout the software design and testing. It also incorporates evidence-based self-management features to facilitate knowledge, skills, and resources for managing blood pressure and medication regimens. We aimed to gather user feedback and interactions with a conversational agent to ensure high usability. As physical activity is an important component of self-management for hypertension control, this work also examines data from wearables and apps to identify physical activity clusters from a variety of exercise metrics (i.e., steps, intensity, distance, calories burned) and evaluates the association between physical activity and health status. While this specific research examines these data in the context of inflammatory bowel diseases, the use of real-world data to identify subgroups with similar activity attributes could also provide insights into how to personalize conversational agents' behavioral coaching strategies and dialogue for other chronic diseases such as hypertension. The overall aims of this dissertation research are:

Aim 1: To assess patient needs and preferences for a conversational agent to support hypertension self-management.

Aim 2: To design and develop, gather feedback, and evaluate the usability of a hypertension self-management conversational agent.

Aim 3: To cluster physical activity phenotypes and assess the association between health status and physical activity from wearable devices, which could be used in future studies to personalize the conversational agent's coaching and information.

To achieve these aims, I first conducted in-depth semi-structured interviews to elicit information and support needs related to using a conversational agent for hypertension medication self-management. Barriers, facilitators, perceptions, and perceived frequency of use were also assessed. Thematic analysis was conducted to identify key content and functionalities to inform the design of the conversational agent. Based on these patient-defined features and users' needs, I then iteratively designed and developed a functional statistical data-driven conversational agent prototype. Usability was evaluated through task-based think aloud methods followed by a usability questionnaire and brief semi-structured interview. Task completion rate, error rate, interaction metrics (i.e., duration, utterances, clicks), and questionnaire scores were calculated. Thematic analysis was conducted to identify the strengths and shortcomings of the prototype. Lastly, I used wearable and app data over a 5-year period to cluster physical activity phenotypes. I used an unsupervised learning approach to generate the clusters and then assessed the relationship between each cluster and self-reported health status. I also identified patients who moved into different physical activity clusters longitudinally and examined changes in their health status. This approach should be validated and then could potentially be used to facilitate contextually aware lifestyle recommendations and tailored coaching strategies within conversational agents.

To date, the optimal design, features, and usability of conversational agents for self-management of hypertension and other chronic diseases are not well known. This research contributes towards a better understanding of the patient perspectives for using a conversational agent to support hypertension self-management. It highlights the unique considerations of using conversational interfaces for managing health, including patient safety, privacy, and integration with other mHealth data. Our findings could also contribute towards the development of a framework that leverages digital phenotyping and conversational agents for just-in-time adaptive interventions to provide meaningful self-management related information to the patient.

Organization of the Dissertation

This dissertation is organized into five chapters. The full methodologies and results of each Aim are summarized in three separate manuscripts (Chapters 2 through 4). In Chapter 2, I assessed patient informational needs and preferences for using a conversational agent for hypertension self-management. In Chapter 3, I designed, developed, and evaluated the usability of the conversational agent prototype. In Chapter 4, I used an unsupervised learning approach to identify clusters of physical activity phenotypes and assessed the relationship among health status and physical activity clusters generated from wearables and apps. The conclusions and directions for future research are summarized in Chapter 5. As much of this dissertation research took place amidst the coronavirus disease 2019 (COVID-19) pandemic, which has transformed the health care and mHealth landscape, additional discussion of the use and impact of conversational agents for COVID-19 are discussed in Chapter 5. Appendices include all study protocols, interview guides, and survey instruments used in this research.

CHAPTER 2: INFORMATION NEEDS AND PREFERENCES FOR A CONVERSATIONAL AGENT FOR HYPERTENSION SELF-MANAGEMENT

Background and Significance

Almost half (46%) of U.S. adults have hypertension[96], making it the most common chronic disease and a leading risk factor for heart disease[97]. While there are multidimensional health system and patient-related factors associated with inadequate blood pressure control, a main contributor is poor self-management related to diet, exercise, sleep, and medication management[98]. Medication self-management is defined as the extent to which a medication is taken as prescribed, including the appropriate dose, frequency, spacing, and safe use over time[99]. Although approximately 75% of U.S. adults take hypertension medications, only half have blood pressures that are adequately controlled[100]. Hypertension is typically asymptomatic, and patients may not perceive immediate benefits from taking medications or adhering to strategies that promote lifestyle changes[12]. Digital health approaches to improve hypertension medication self-management have been investigated using mobile applications (apps), short messaging services (SMS), and devices that connect to apps, such as Bluetooth pill boxes[27, 38, 42]. Prior research has demonstrated improvements in medication self-management primarily through informational, behavioral, and motivational approaches, such as education, tracking, reminders, and social support[27, 38, 42, 101-103]. However, many digital health approaches often fall short due to suboptimal adherence towards the technologies and limited patient engagement[42, 104]. This may be due, in part, to limited personalization of digital solutions or lack of user motivation to interact with the intervention[104-106].

Emerging digital health technologies, such as conversational agents, have the potential to communicate with patients and serve as effective self-management tools for chronic conditions[44, 46, 66, 70, 71, 107, 108]. Conversational agents, also known as chatbots, are systems that can communicate in natural language through text or voice[45]. Very few studies have assessed the use of chatbots for hypertension self-management[109, 110]. Persell et al (2020) evaluated the effectiveness of a home blood pressure monitor plus text-based chatbot, which provided encouragement for blood pressure tracking, medication adherence check-ins, and coaching for barriers to adherence[109]. Although no differences were found between groups for mean blood pressure or medication adherence at six months, self-confidence in controlling blood pressure was significantly higher in the chatbot intervention group[109]. Migneault et al (2012) assessed the use of a culturally adapted, voice-based system, which provided coaching on medication adherence, physical activity, and diet in patients with hypertension. No differences were found in medication adherence or blood pressure between groups over the one year study period, but there were improvements in diet quality and energy expenditure[110]. These early findings suggest the optimal design, features, and preferences for using conversational agents for hypertension self-management are not well known, and there may be potential design improvements to facilitate blood pressure control.

Objectives

The objectives of this aim were to elicit information needs and perceptions towards using a chatbot to support hypertension medication self-management as an initial phase of the user-centered design process towards developing a chatbot prototype.

Methods

User-Centered Design Framework

User-centered design, developed by Norman and Draper (1986), is an iterative multi-stage process in which user feedback is involved throughout each stage to ensure high usability[111]. The Integrate, Design, Assess, and Share (IDEAS) framework by Mummah et al (2016) builds on user-centered design by incorporating health behavior theory into digital health interventions and highlights the importance of evidence-based implementation strategies[112]. The integrate phase focuses on gathering qualitative insights from users and defining behavioral goals within the intervention[112]. This is followed by the design phase which incorporates users' insights into iterative prototypes and user feedback through usability testing, surveys, or interviews. The assess phase evaluates the efficacy of intervention through pilot studies or randomized controlled trials. Lastly, the sharing phase involves disseminating findings to advance research and practice.

The IDEAS framework has been leveraged throughout several digital health interventions[113-115]. The framework has also been used in the development of a conversational agent to reduce alcohol use, in which focus groups were conducted to assess user needs, followed by ideation and iterative prototyping to incorporate user input and behavior change theories[68, 69]. In this study, we focus on the initial integrate phase to gather user needs and insights towards using a conversational agent to facilitate hypertension medication self-management.

Study Design

We used a convergent mixed methods design[116] to assess patient information needs and perceptions towards a text-based hypertension medication self-management chatbot, which combined qualitative and quantitative data collected from in-depth semi-structured interviews and self-administered questionnaires. Because several sociodemographic characteristics are associated with access and use of technology (e.g., age, education)[117, 118], a mixed methods approach was used to generate a more comprehensive understanding of patients' perceptions across sociodemographic characteristics to optimize the design of the chatbot. This study was reviewed and considered exempt by the University of North Carolina at Chapel Hill Office of Human Research Ethics Institutional Review Board.

Study Setting and Participants

Participants were adults (18 and older) who self-reported having hypertension, took at least one hypertension medication, and resided in Chapel Hill or surrounding areas in North Carolina. Eligible participants had to speak English, take their own medications without assistance, have the ability to provide informed consent, attend an in-person interview, and own a smartphone or tablet. Purposive sampling was used to select 10-15 adults based on age, race, gender, education, and number of prescribed medications. This sample size was chosen because prior research has shown thematic saturation generally occurs within the first twelve interviews[119]. Interviews were conducted until data saturation was reached, which was defined as when no new themes emerged from the data[120].

Recruitment

Participants were recruited from websites, e-mail list-servs, and flyers posted in clinics, hospital waiting areas, and community locations around Chapel Hill, North Carolina and the surrounding area. Recruitment materials contained a link to an electronic screening questionnaire to assess eligibility. Then participants were sampled to vary representation among clinical and sociodemographic characteristics. The aim was to have at least: five adults who were 65+ years, five of minority race, five males, five without a college degree, and five who were taking at least three medications.

Interview Guide Development

Our multidisciplinary research team developed an interview guide to assess information needs and perceptions of using a chatbot for hypertension self-management. The interview guide was informed by constructs from the Information–Motivation–Behavioral Skills Model[121] and the Unified Theory of Acceptance and Use of Technology (UTAUT)[122]. The interview guide contained questions related to information needs such as “What types of information or resources would be helpful for you to keep track of taking your medications?” and “What type of support may be helpful for you to keep your blood pressure under good control?” The UTAUT has also been extensively used to understand acceptance and behavioral intention to use health technologies[123-125]. Questions in our interview guide related to these constructs included: “How do you think a chatbot could help you take or refill your medications?” and “How often would you want to interact with a chatbot?” Follow-up questions to probe were asked as needed to generate additional insights. See Appendix A for the full interview guide and protocol.

Information-Motivation-Behavioral Skills Model

The Information-Motivation-Behavioral Skills Model describes how health-related information, motivation, and behavioral skills are determinants of health behaviors[121]. This model has been widely used in understanding self-management behaviors and needs[27, 38, 42, 101-103, 126, 127]. Information includes materials used to inform or make health decisions (i.e., educational resources about hypertension, medication dosing, side effects, appointment scheduling, care team contacts)[27, 37, 38, 102, 103]. Motivation involves personal or social support for enactment of the health-related behavior and is often influenced by confidence and intrinsic value placed on adhering to the treatment plan[42, 102, 103]. Thus, supporting self-efficacy and providing empathy may facilitate the ability to manage hypertension[128-131]. Behavioral skills are the actual and perceived ability to enact a health behavior, and digital strategies have focused on tracking and reminders for medications, refills, and blood pressure[27, 37, 38, 42, 102, 103]. Therefore, when individuals are well-informed about their health condition and medications, motivated to adhere to their regimen, and have the necessary behavioral skills for effective management, they are more likely to initiate and sustain treatment adherence[121].

Unified Theory of Acceptance and Use of Technology

The Unified Theory of Acceptance and Use of Technology (UTAUT) has seven constructs that have a direct role in determining the behavioral intention to use a technology and usage behavior: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price, and habit[122]. Gender, age, and experience are moderators for each construct and behavioral intention. Within the context of digital health, performance expectancy is degree to which patients believe that using the technology will enable them to improve their health condition[124, 132]. Evidence suggests most individuals believe health technologies may be useful to monitor their health status and treatment plan[123, 124, 133]. Effort expectancy refers to the ease of use associated with using the technology and has been identified as one of the most important factors directly influencing digital health adoption and acceptance[123, 124]. Social influence is the degree to which patients perceive they should use the technology based on the importance of other people, and patients are highly influenced by the opinions of their care team[133]. Facilitating conditions relates to beliefs that the technical and organization infrastructure supports the use of the technology[124, 132]. Older and less experienced patients typically require the most assistance[124, 132]. Hedonic motivation is the pleasure obtained from using a technology with younger patients exhibiting a greater tendency to seek out novel technologies[122]. Price is also a predictor of behavioral intention to use a technology, and evidence suggests that the popularity of SMS is due to low pricing[122]. Habit is the extent to which patients tend to perform a task automatically, and the degree of habit is stronger for those with more experience with the technology[122]. Collectively, these factors show that digital health technology acceptance (usage behavior) is largely dependent on a patient’s behavioral intention to use it, habits, and the conditions that facilitate use.

Questionnaire Development

The questionnaire contained questions about demographics, clinical characteristics, experience with using technology, and additional validated questionnaires on the topics of health literacy[134], medication self-efficacy[135], and barriers to medication adherence[136] (see Appendix A). The 3-item Brief Health Literacy Screener by Chew et al (2004) was administered to assess health literacy[134]. Total scores range from 3 to 15, and any response greater than 3 for any question indicates inadequate health literacy[137]. The Patient-Reported Outcomes Measurement Information System (PROMIS) Self-efficacy for Managing Medications and Treatments short form 8a (2016) was used to assess confidence in managing medication schedules and treatments[135]. The instrument was scored using the PROMIS HealthMeasures Scoring Service, where raw scores were converted into T-scores with a mean of 50 (SD=10) with higher scores representing greater self-efficacy. The Adherence Starts with Knowledge 12 (ASK-12) by Matza et al (2009) assessed barriers to medication and treatment adherence in three domains: inconvenience or forgetfulness, treatment beliefs, and behaviors[136]. Total scores range from 12 to 60 with higher scores representing greater barriers to adherence. Prior approval was obtained for the use of 3-item health literacy measure and ASK-12[134, 136].

Interviews

The interview guide and questionnaire were initially pilot tested with several members of the study team and then a trained interviewer (AG) conducted in-person, semi-structured interviews using the interview guide. First, participants were consented and completed the study questionnaire. Next, participants were asked about their current self-management behaviors and information needs to manage their blood pressure and to support their medication regimens. Then, to introduce what a chatbot might look like, participants were shown a short video of a commercial text-based health chatbot “Florence”[138, 139], which provides medication reminders and health tracking functions. Participants were also asked about their perceptions towards using a chatbot to manage their blood pressure and medications. Additionally, perceived barriers and facilitators for using a chatbot were elicited. Each interview lasted approximately 60 minutes, was audio-recorded, and participants were provided with a $25 gift card.

Analysis

Interviews were transcribed verbatim, and imported into NVivo qualitative data analysis software[140]. Participant narratives within the transcriptions were analyzed using applied thematic analysis[141]. First, a codebook of structural codes was developed based on the initial topics from the interview guide prior to analysis. Two independent reviewers (AG, ZX) applied these structural codes to segment participant narratives by topic. Discrepancies in coding were adjudicated by a third reviewer (SM) when necessary. Transcripts were initially double-coded until Cohen's kappa[142] of 0.8 was reached, which was after three transcripts. After double-coding, the rest of the transcripts were equally distributed and single-coded by the reviewers, and discussion occurred after every 2-3 transcripts. Next, each reviewer inductively identified and applied thematic content codes in each structural coding report with each report containing a topical area across all participant narratives (e.g., perceptions, barriers, facilitators, etc.). Structural coding reports were also initially double-coded and Cohen's kappa was assessed again as above, which was after three reports. The remaining reports were then equally distributed and single-coded among reviewers (AG, ZX), and reviewers discussed whether new or additional content codes should be added after each report. Throughout the entire analytic process, the codebook was iteratively revised based on disagreements and emerging topics, and transcripts were recoded as needed. Lastly, reviewers met to organize the content codes thematically to describe the major themes, subthemes, and illustrative quotes within the themes. Quantitative data collected from the questionnaire were summarized using descriptive statistics. Sociodemographic data were integrated with the perceptions and perceived usage themes to better understand attitudes toward using a chatbot.

Results

Sample Characteristics

Thematic saturation was met after interviewing 15 participants. The average age was 59 years, eight (53%) were female, ten (66%) were White, and nine (60%) had at least a college education (Table 2.1). Nine (60%) had hypertension for at least five years, and nine (60%) were "very or completely confident" their blood pressure was under control. On average, participants had three comorbidities and were taking six medications. The majority of participants (87%) had adequate health literacy, were above the U.S. population average for medication self-efficacy (52.3), and the greatest barrier to adherence was behavior (i.e., "not had a medication with you when it was time to take it"). Only 20% of participants reported using a chatbot before.

Table 2.1 Sample characteristics

Characteristics	**n (%)**
Age (mean=59, SD=11)	
45 – 54 years	6 (40)
55 – 64 years	4 (27)
65+ years	5 (33)
Gender	
Female	8 (53)
Male	7 (47)
Race	
White or Caucasian	10 (66)
Black or African American	4 (27)
Other	1 (7)
Ethnicity	
Not Latino/Latina	14 (93)
Latino/Latina	1 (7)

Education Less than college College graduate or more	 6 (40) 9 (60)
Household Income $20,000 – $34,999 $35,000 – $49,999 $50,000 – $74,999 $75,000 or more	 2 (13) 4 (27) 5 (33) 4 (27)
Comorbidities, mean (SD)	3 (1)
Years with Hypertension Less than 1 year 1 – 2 years 3 – 5 years 5 or more years	 1 (7) 3 (20) 2 (13) 9 (60)
Number of Prescription Medications (mean=6, SD=4) 1-3 4-6 7+	 6 (40) 4 (27) 5 (33)
Confidence Blood Pressure is Under Control Somewhat confident Very confident Completely confident	 6 (40) 5 (33) 4 (27)
Internet Use Once a day More than once a day	 1 (7) 14 (93)
Device Use Smartphone Tablet Computer	 15 (100) 10 (66) 13 (87)
Ever Used a Chatbot Yes No/Don't Know	 3 (20) 12 (80)
Health Literacy Level[134] Adequate Inadequate	 13 (87) 2 (13)
Medication Self-Efficacy[135], mean (SE)	52.3 (4.0)
Barriers to Adherence[136], mean (SD) Behaviors Treatment Beliefs Inconvenience/Forgetfulness Total Score	 8.0 (3.9) 7.7 (2.0) 6.1 (2.8) 21.8 (6.1)

Information Needs for a Hypertension Medication Self-Management Chatbot

Qualitative analysis identified four domains (medications, refills, communication with the care team, and healthy lifestyles), which comprised ten themes for information and support needs for hypertension medication self-management (Table 2.2). Medication information needs included: having a list of current and past medications with frequency and dosage, the ability to set reminders to take medications, and information about side effects, medication interactions, and similar medications. The majority of participants wanted reminders only on weekends, holidays, or when there might be variance from their normal routine. For managing refills, participants were primarily interested in reminders to order or pick-up medications and the ability to view the number of refills left, date of next available refill, and expiration date. Some wanted a chatbot to integrate with their pharmacy to automatically order refills. Most desired to communicate with their care team by sharing their health data (e.g., blood pressure, weight, physical activity) and to be able to schedule appointments. For healthy lifestyles, the majority were interested in tracking health-related metrics and receiving encouragement based on these data from the chatbot. Several described how a chatbot could provide feedback on results after a clinic visit, and many felt it would be necessary to integrate the chatbot with existing apps, specifically MyChart and Fitbit. Several expressed the need for accountability to keep their blood pressure under control.

Table 2.2 Self-management needs and desired features of a chatbot for hypertension self-management

Themes for User Needs	**Representative Selected Quotes**
Medications List of current and past medications	“I can never remember the name [of the medication]…I wonder if that may be able to hold a history on your medications.”

	"It would be lovely to load in my medications because I have an extensive list."
Reminders	"If it could prompt me only on days where there's a high probability that I forgot [to take the medications], like weekends and holidays..." "I might only need my reminder once every three or four days to make sure that I'm where I need to be with my medications…"
Information about medications and side effects	"I'll use various websites, sometimes WebMD or Mayo Clinic, to see what they're giving me and what side effects I can look for…I always ask how it interacts with my other medicine, and if I'm allowed to adjust the times according to what works best for me."
Refills Reminders to order and pick up medications	"If that system worked the way you really wanted it to, you would put in your medication, milligrams, frequency, and how often it refills. Then, it'll prompt you and say, 'It's time to refill your medicine.'" "When [the chatbot] says, 'Okay, it's time for a refill.' [I'd say] 'Can you request the refill without me going through the extra steps?'"
Number of refills left and dates of next refill and expiration	"It might be useful to see the refill date so I would know without having to count how many I have left." "It'd be nice for my health professionals to help keep up when the prescriptions expire because I run into that a lot...Send a reminder or set it up so that it lets me know, 'Okay, you have one month of this left.'"
Communication with care team Sharing health information	"It'd be really cool if there was some way when I take my blood pressure I could get it into my medical records…If the chatbot was something where I could put in my readings, when I go to my doctor, [I could] bring it in or go through it with the nurse during our quarterly call." "I got one doctor that's my primary doctor, and I've got two more doctors which are my cancer doctors. They ask me about each other. That right there'd be able to help me communicate; get them on the same level."
Appointment scheduling and reminders	"It would be helpful if the chatbot could remind me a month before, 'You need to schedule your appointment.' Or, 'It's time for your physical.'" "I would much rather schedule my appointments through [the chatbot] than having to call. I don't like calling because they want me to repeat my entire life story and give all my information to a person I don't know."

Healthy lifestyles Health tracking and compatibility with apps	"A lot of times my weight and blood pressure are tracked through my MyChart app…If all of that can be fed into the chatbot, it would be a better tracker because it would have a more rounded view." "MyFitnessPal has weight. Fitbit can do weight, but that's a premium feature…An application like this probably should tie into both. I'd be interested in letting them connect."
Feedback and encouragement	"Apple has the Health app, and it tracks everything…it actually reads from MyChart app now. It collects, but it doesn't communicate. At best, it tells you [that] you have new information in your chart…[The chatbot is] a place where you can actually have almost a dialogue." "If there was some type of way [the chatbot] was able to check what my blood pressure was at the time it's elevated, then it would [say], 'It's time to take a break. Maybe you should go for a walk."
Accountability	"I think it would be great because it's telling you, 'Do it'…With MyFitnessPal, I'm just looking to see how many steps I did. With that one, it's going to probably prompt you for more things." "It can probably track my last time taking my weight…It's on you at all times, so that's what I like about that."

Perceptions and Perceived Use of a Chatbot

Perceptions for using chatbots were categorized into three key themes: similarities to existing apps, curiosity about chatbots, and chatbots being humanlike (Table 2.3). The majority of participants compared chatbots to apps they currently use to track or manage their health such as MyChart, Fitbit, Apple Health, and health insurance apps. Many were curious to know if a chatbot already existed, and those who expressed curiosity were slightly younger on average than those who did not (56 vs. 63 years). However, some older adults conveyed interest in using new technologies specifically because they were older and wanted to keep up with emerging technologies. Most of the participants who were curious about using a chatbot were taking fewer medications on average compared to those who did not (4 vs. 7 medications). The majority who were taking several medications felt they had already established a routine and would not rely on a chatbot for a reminder. Several also felt chatbots seemed humanlike and compared them to talking with a friend or health coach. Those who perceived them as humanlike were younger on average (51 vs. 64 years) and taking fewer medications (3 vs. 6 medications) compared to those who did not perceive this. Overall, perceptions of chatbots may vary based on age and number of prescribed medications taken.

Perceived frequency of use of a chatbot was grouped into three categories based on analysis: daily, weekly to monthly, and rarely to never (Table 2.3). Several described how the amount of interaction would depend on utility of information provided. Similar to patient perceptions, those taking fewer medications were interested in more frequent interactions, especially for non-medication purposes such as blood pressure tracking. Many who were diagnosed with hypertension within the past five years did not feel very confident that their blood pressure was controlled. Several of these participants described being open to different approaches that might help. Overall, perceived frequency of interaction differed across characteristics including number of medications, time since diagnosis, and the level of confidence that blood pressure was under control.

Table 2.3 Perceptions and perceived frequency of use of a chatbot

Themes for Perceptions	**Representative Selected Quotes**
Similarities to existing apps	“I go through MyChart now to do most of [the appointment scheduling], and I guess that’s kind of like a chatbot.” “It reminded me of the United Health app. That’s pretty neat.”

Curiosity about chatbots	"Is this being used at all, or are we totally in testing mode for this thing?…It's pretty fascinating stuff." "I like that – do they have it already?...I don't want to miss it, and not be able to have something like that."
Humanlike	"It was like you were just texting a friend, so it looked friendly and inviting." "You would think really that you were talking to a person in a lot of ways."
Themes for Perceived Frequency of Use	**Representative Selected Quotes**
Daily	"I would probably use it on a daily basis, almost. It's right there on the phone…I'd love to try it." "I wouldn't mind [using it] every day. I have a lot of apps I interact with every day."
Weekly to Monthly	"It would be useful if I could decide how much stuff I'm getting…For the health tips, maybe once or twice a week…" "Every few weeks would be fine unless I really had some follow-up stuff to do or if I was having a problem"
Rarely to Never	"I probably wouldn't use it…I would find it unnecessary because I think I have under control what I can control." "I'm sure there are folks who take advantage of things like that. Maybe at some point I would, but right now, no…If things start getting too hectic, [I need to] slow down..."

Barriers and Facilitators of Using a Chatbot

Four main themes were identified for barriers and three themes were identified for facilitators of using a chatbot to help manage medication regimens and blood pressure (Table 2.4). Barriers included cell phone issues, fears that the chatbot would provide excessive or unhelpful information, make demands, or invade one's privacy. Several participants were concerned that a cell phone screen would be too small or that keeping track of their cell phone regularly would be difficult. Although some felt their blood pressure and medication routine were already under control, a few mentioned non-blood pressure use cases for self-tracking, such as cancer self-management or managing a family member's health. Several stated that they specifically did not want a chatbot to tell them what to eat or to lose weight. In regards to privacy issues, some participants referenced Amazon Alexa and Google Assistant and were worried about a chatbot listening to their conversations or sharing their information with other companies.

Key facilitators for using a chatbot for medication self-management included customizability, convenience, and being unobtrusive. Nearly all wanted to personalize the chatbot, especially the frequency of reminders and tips. Among those who reported using a chatbot before, all of them discussed the importance of tailoring the amount and type of information to make the interactions useful. Most liked the convenience of a chatbot being accessible on their cell phone or having all of their health information in one location. However, many also did not want the chatbot to interrupt day-to-day activities.

Table 2.4 Barriers and facilitators of a chatbot for hypertension self-management

Themes for Barriers	**Representative Selected Quotes**
Cell phone issues	"I can go to MyChart, but I normally do that on the big computer. It's just kind of aggravating on my smartphone. I don't know how that chatbot might be."
Too much information or not useful information	"It just felt like it was annoying, had too much information, and I didn't want look at it cause it's too many things to go through…" "To open the computer or iPad to get additional health information would be unnecessary because I'm pretty conscientious with my medicine."

Making demands	“It’s either going to be a good conversation with the chatbot or it could get a little lippy if I put some weight on, in which we would reduce the chatbot usage to once a week.” “Telling me, ‘Don’t eat that burrito. There’s too much salt.’…I don’t want to go to my iPhone to ask if I can eat my burrito.”
Invasion of privacy	“I’m not going see a message on my TV that [says] ‘Did you take your medicine?’ or Alexa’s not going to tell me, ‘You better check your phone.’ I get creeped out when technology is intrusive.” “I have one of those Google speakers at home. I unplug it when I’m home because sometimes I’ve had a conversation and it picks it up. The next thing you know I’m getting advertisements…As long as it wasn’t intrusive like that.”
Themes for Facilitators	**Representative Selected Quotes**
Customizability	“If you could check some boxes of things you like and don’t want…You could check: I want tips daily, weekly, monthly, no tips, or I want reminders every day for checking my blood pressure.” “I imagine she will pop up on my phone and say ‘Take your meds’, ‘It's time for a refill’, or ‘You don't have refills left’…There should be some flexibility in scheduling it like there is with your calendar.”
Convenience	“I think it’d be helpful with managing my health…It’d be nice to have it all in one place on my phone to use it whenever needed.” “It’s more like a health coach instead of me waiting for somebody to call me…I miss the call, call back, and then I have to go through many things to get to the person. I think that would be a little bit more helpful because I’m not trying to run someone down.”
Unobtrusiveness	“I prefer things to be as minimal and as automatic as possible…That’d be great if it’s streamlined and invisible.” “I would be interested in some kind of way of recording stuff as long as it wasn’t so time consuming.”

Discussion

Principal Findings

Chatbot-related needs for medication self-management in patients with hypertension were consistent with patient needs identified in prior research focused on medication adherence and self-management[27, 38, 42, 101-103]. Our research extends these perspectives by providing additional understanding and nuance around leveraging chatbots for hypertension self-management. Overall, participants had generally positive attitudes towards medication self-management interventions delivered via chatbots. While most had not previously used chatbots, almost all perceived the conversational nature of chatbots to be potentially helpful for various self-management tasks and encouragement for healthy lifestyles. Using the chatbot to assist with tasks such as tracking medications, refills, blood pressures, or communicating with care team members were felt to be particularly impactful. Many believed chatbots would be valuable if tailored and compatible with patient portals, pharmacy apps, or health tracking apps.

There may also be differences in the perceptions and possible use of chatbots based on sociodemographics. Patients who were younger and taking fewer medications were the most curious about using a chatbot. The Unified Theory of Acceptance and Use of Technology demonstrates how age and experience with technology are moderators of one's acceptance and behavioral intention to use a technology[122]. Thus, it is likely that younger patients, who are typically more tech savvy and taking fewer medications than older patients, may be more interested in using a chatbot. Characteristics varied additionally with participants' perspectives on how they often might desire to interact with chatbots. Despite the enthusiasm for using chatbots for hypertension self-management, participants expressed several concerns with chatbots providing too much information, messages about lifestyle modifications being demanding, invading their privacy, or usability issues with interacting with chatbots on a cell phone.

Implications for Health Care and Research

Our study revealed several design recommendations and implications for hypertension and medication-related conversational interfaces and user experience. Similar to prior research findings[143], the ability for individuals to personalize nearly all aspects of the chatbot is important, including tailoring content, frequency of receipt information and reminders, tone of language, and type of feedback. Several participants in our study mimicked the type of feedback and conversation they would like to have with the chatbot and this differed across participants. This is consistent with other studies, which found that individuals may have varied needs for different types of support depending on sociodemographic characteristics, health status, and cultural factors[44, 77]. Some patients may prefer a more active coaching style based on their established routines, self-efficacy, confidence, or duration of hypertension, while others may feel it is too intrusive. Because patients with hypertension often have other comorbidities, tailoring information for a broader base of conditions or based on a recent diagnosis may be necessary to better meet patients' informational needs within that context.

To maximize utility, it is important to consider whether tethering chatbots to other applications to integrate health information from portals, pharmacies, or other health apps may improve the user experience or self-care tasks. For example, chatbots could provide an enhanced interaction by contextualizing data, such as number of daily steps taken or a lab result, into actionable health behavior changes. A survey of physicians found overall positive perceptions towards chatbots being able to support and motivate patients[144], though workflow barriers and facilitators of integration within clinical settings have not been well-studied. Rich linguistic data generated from chatbots could also provide valuable information for patient health status to the care team. Voice biomarkers and patterns of interactions with smartphones have previously revealed changes in pulmonary hypertension and cognitive functioning over time[145, 146]. However, using systems with unconstrained natural language input capabilities comes at greater risk for potential errors in natural language understanding[44]. Careful considerations must be made about the safety and quality of information provided in these systems[88, 147]. Additional considerations for incorporating linguistic data from chatbots, mapping these data to existing terminologies, and interoperability within health systems are also needed for chatbots to be integrated into clinical care. As with incorporating any type of patient-generated health data into care settings[148], the relevancy and interpretability of health professionals should also be assessed.

Limitations

Our purposeful sample was limited to adults with a smartphone or tablet from a single geographic location in the Southeast, so these findings may not reflect the perceptions among all adults with hypertension. As purposeful sampling is a non-probabilistic sampling method used in qualitative studies, we were unable to control for the potential influence of confounding variables or differences in perceptions based on individual sociodemographic characteristics. Three participants (20%) reported prior experience with a chatbot which might affect their perceptions and perceived use.

Although all participants watched a video of the same example of a chatbot, those with no prior experience may have found it more challenging to envision interactions with a chatbot. Moreover, those who agreed to participate may have had stronger inclinations towards using new technologies and may not be representative of all patient perspectives. Our sample also had high levels of adequate health literacy and above average medication self-efficacy in comparison to the U.S. adult population, which may also limit generalizability.

Conclusion

Given the growing burden and national focus on hypertension control in the U.S., novel self-management and medication adherence tools could help improve blood pressure control, which could be impactful both for patients and health systems[149, 150]. Although most participants (80%) had never used a chatbot, the majority showed an interest in using a chatbot to help track their medications, refills, blood pressures, or communicate with their care team. Our findings contribute to a better an understanding of user needs and perceptions towards using a chatbot for hypertension self-management across individual characteristics, such as age, number of medications taken, years of diagnosis, and level of confidence about blood pressures being under control. Being mindful of innate user differences can facilitate the design and development of user-centered, personalized chatbot interventions. While the use of chatbots for self-management is still an emerging area, chatbots have the potential to not only provide evidence-based resources, but to also actively engage patients through conversational dialogues about their health information and goals. This research can be used to inform the future design and functionalities of conversational interventions to support hypertension medication self-management. Additional investigation is needed to assess the usability, optimal timing and type of support, appropriate dialogues and interactions, and the privacy implications of chatbots.

CHAPTER 3: DESIGN, DEVELOPMENT, AND USABILITY OF A HYPERTENSION SELF-MANAGEMENT CHATBOT

Background and Significance

Chatbots, which are dialogue systems that mimic human chat characteristics, are increasingly used for business, e-commerce, or education as automated online assistants to provide customer service support or route users to a human[87]. Chatbots in health care have been less adopted, but early evidence suggests their potential to support patients through self-monitoring, psychotherapy support, and educational resources[44]. At this nascent phase, research for health care chatbots has largely focused on mental health conditions[44, 108]. For example, text-based chatbots have demonstrated reduced depression and anxiety symptoms, reduced stress, and improved flourishing[70, 71, 89]. The initial focus on mental health may be due to the early psychology use cases for chatbots. The first well-established chatbot, ELIZA, simulated a Rogerian psychotherapist[47]. Rogerian psychology led to cognitive behavioral therapy and motivational interviewing, which are client-centered counseling styles for eliciting coping skills and behavior change[48]. As self-management skills and behavior change are important components in the prevention and management of chronic diseases, diverse health care use cases for chatbots have begun to emerge for diseases such as diabetes, obesity, and hypertension[72, 109, 110, 151].

Hypertension is the most common chronic disease and leading risk factor for heart disease in the U.S.[96, 97]. To date, two randomized controlled trials have evaluated the efficacy of hypertension-related dialogue systems[109, 110]. These systems had improvements in self-confidence in controlling blood pressure, diet quality, and energy expenditure[109, 110]. However, there were no differences in blood pressure control or medication adherence between intervention and control groups[109, 110]. These early findings may indicate that the optimal design and user interactions of chatbots to support hypertension self-management are not well delineated. Involving patients throughout the design process and ensuring technologies are usable have illustrated positive improvements in patient empowerment and safety[152]. Usability refers to the ability for users to achieve their goals with efficiency, effectiveness, and satisfaction[63], and usability challenges present a major barrier to health information technology (HIT) adoption[153]. Usability issues may lead to user frustration, inability to achieve goals, and prevent long-term engagement needed to sustain ongoing self-management behaviors.

Although there is limited research on chatbots for self-management, chatbots have generally been reported as acceptable and usable by patients[72, 73, 89, 154, 155]. Patients have also described chatbots as friendly, useful, and able to provide enjoyable interactions[89, 155, 156]. Older adults and patients with low health and computer literacy have expressed positive attitudes and found chatbots easy to use[73, 154, 155]. Nevertheless, most chatbots are still limited by their inability to hold meaningful conversations and personalized interactions[87]. To assess their usability, chatbots have been evaluated on a number of technical performance and use metrics, such as dialogue efficiency, response generation, response understanding, speed, error management, task completion rate, esthetics, realism, and satisfaction[157]. These qualitative and quantitative metrics offer methods for understanding and improving the usability of chatbots to better meet patients' needs. As patient-facing technologies are less utilized by patients of lower socioeconomic status and older adults[117], it is critical to involve them during the course of design to prevent technology-driven disparities. In this study, we sought to leverage a user-centered design process to design and develop a chatbot focused on hypertension self-management. We then evaluated the usability to better understand how users interact with chatbots to facilitate self-management tasks for hypertension.

Objectives

The objectives of this study were to design, develop, and evaluate the usability of a hypertension self-management chatbot, called "Medicagent." Additionally, the feedback and insights from the usability testing will inform optimizations of Medicagent to prepare the chatbot for future pilot testing.

Methods

Chatbot Design

Throughout the design process, we used the Integrate, Design, Assess, and Share (IDEAS) framework, which leverages user-centered design and theory-driven behavioral strategies[111, 112]. For health behavior theories, we used the Information-Motivational-Behavioral-Skills model which illustrates how information, motivation, and behavioral skills are determinants of health behaviors[121]. We sought for users to be informed through evidence-based resources, motivated through communication and support from their care team, and have behavioral skills facilitated by reminders to enact medication self-management tasks. In addition, we leveraged the model of medication self-management to incorporate the steps involved in successful management of medication regimens (i.e., filling, understanding, organizing, taking, monitoring, and sustaining medications)[99]. Specific features of Medicagent were informed by a previously conducted qualitative study that determined informational needs and preferences on the topics of medications, refills, communication with the care team, and healthy resources[158] (see Chapter 2). Medicagent promotes blood pressure control and medication adherence by allowing users to set reminders to take or refill medications, track blood pressure measurements, provide encouragement, and share health data with the care team. Resources were also available for healthy recipes, fitness tips, and stress management techniques. Iterative low fidelity prototypes were created through sketches and conversational flow maps. Designs were iteratively discussed and modified with members of the research team.

<u>Chatbot Development</u>

The system architecture contained three primary components: user interface, chat engine, and database[159] (Figure 3.1). The user interface included a chat dialogue window that could be opened or closed by the user and accessed through a computer or mobile device. Google Cloud Dialogflow natural language understanding platform was used as the chat engine[50]. Dialogue fulfillment tasks were handled using Google Cloud Functions, and Google Cloud Firestore was used as a database. The following steps take the user’s utterance or input through the system until a response is returned to the user (Figure 3.1):

1. The user types an utterance or phrase.
2. Dialogflow Messenger Integration sends the user utterance to Dialogflow

Application Programming Interface (API) Service.

3. Dialogflow API Service matches the utterance to an intent and extracts parameters. It sends a webhook request message to the webhook service, Cloud Functions. The request message contains information about the intent, action, and extracted parameters.

4. Cloud Functions performs actions as necessary, such as retrieving and returning information from Cloud Firestore.

5. Cloud Functions sends a webhook response message back to Dialogflow API Service which contains the response to send to the user. Webhooks could also be used to retrieve, process, or return information from other APIs.

6. Dialogflow sends the response back to the user.

7. The response is returned to the user.

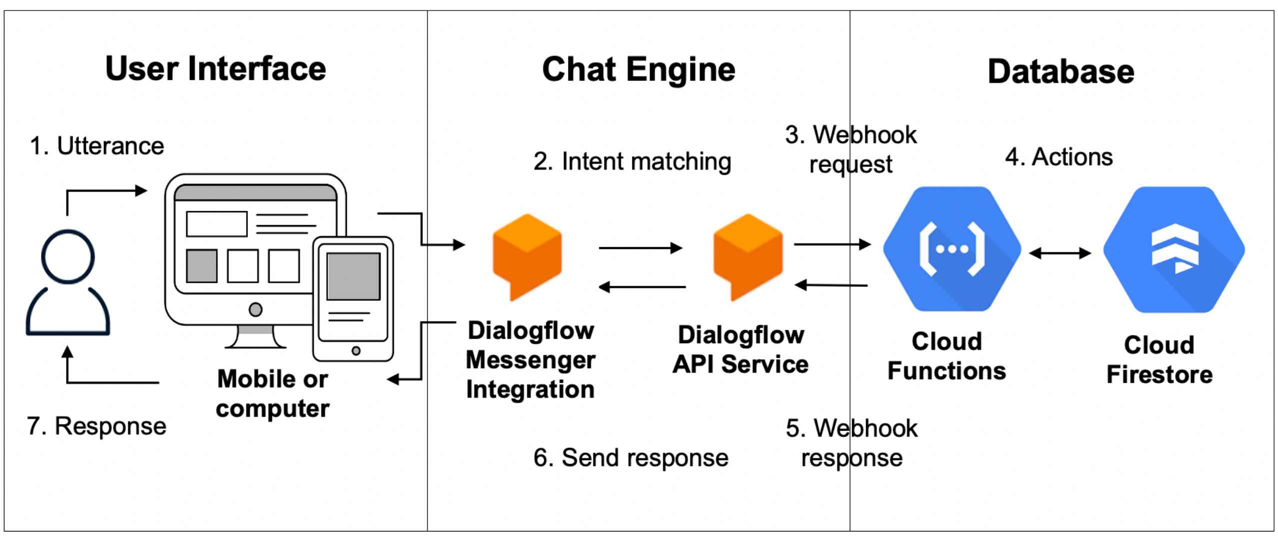
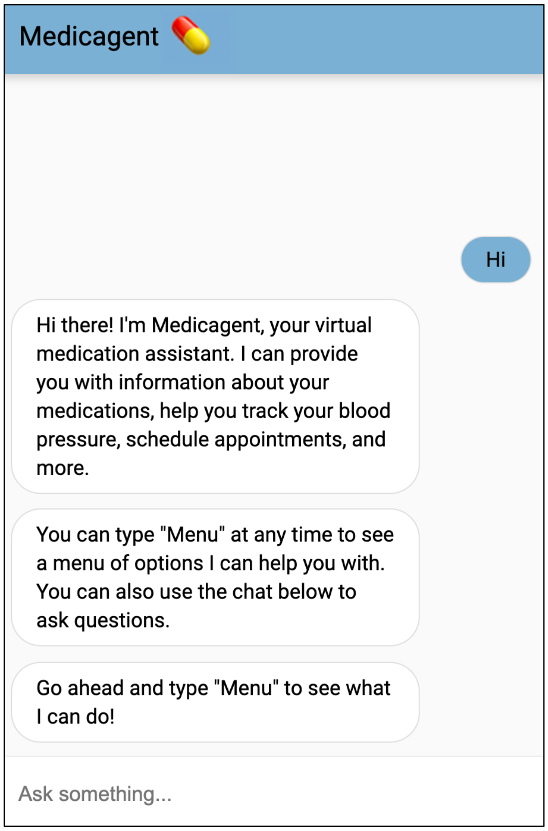
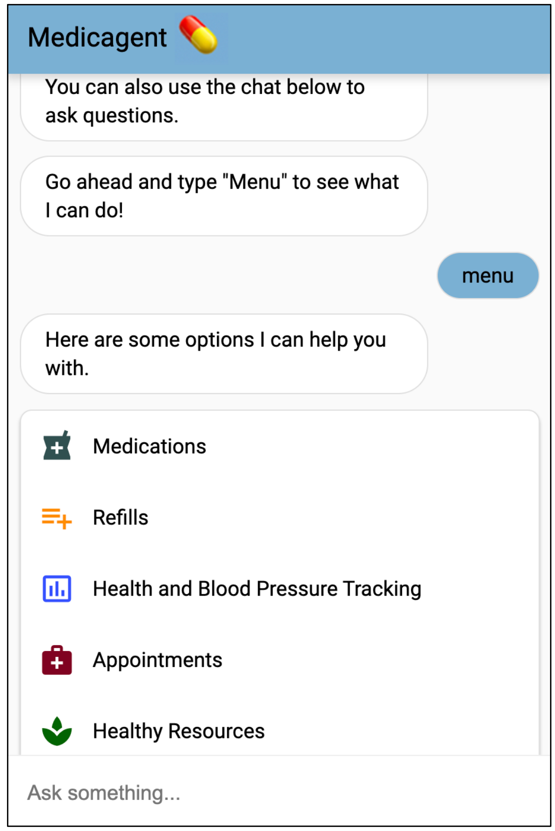


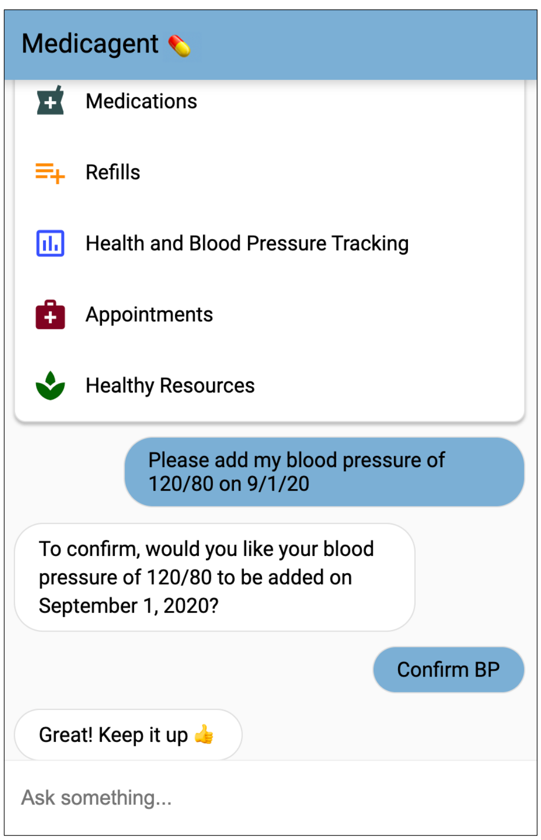
Figure 3.1 System architecture

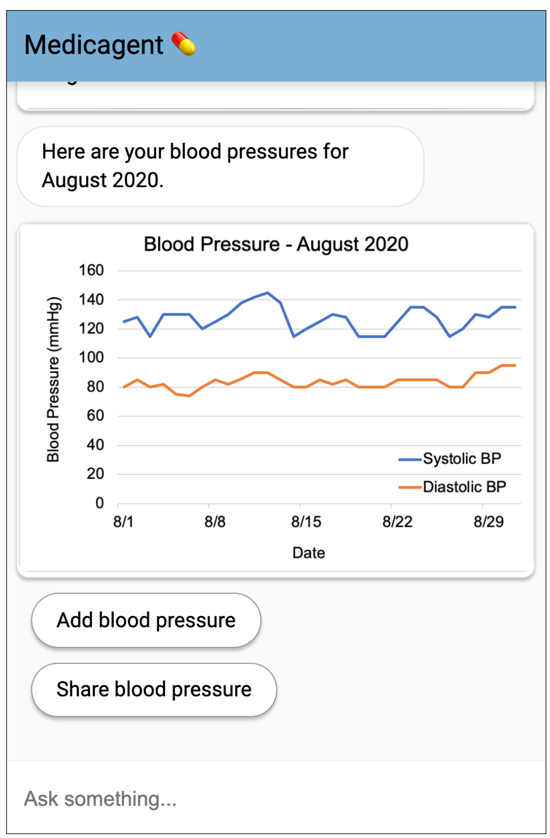
User Interface

The web interface is compatible with most browsers (Chrome, Edge, Firefox, Safari) and responsive on both computer and mobile devices (iOS and Android). When used on a computer (desktop or laptop), the chat dialogue window appears in the lower right side of the screen. On a mobile device, the chat dialogue window fills the entire screen. We aimed for to inclusivity and accessibility for patients of diverse abilities. Web content accessibility guidelines were used such as large font size, color contrast, keyboard navigation, small blocks of text, and appropriate language for low literacy or numeracy users[160]. Medicagent also included visual affordances such as icons, emojis, graphics, and buttons, which are typically familiar to those who use social media or text messaging. The conversation could be driven by the system or the user. A list of synthetic medications, refills, and blood pressure values was preloaded into the database for the initial prototype. Content and images were derived from the U.S. National Library of Medicine's Pillbox[161] and MedlinePlus[162]. The user initiates the conversation with Medicagent by typing a variation of "hi" (top left in Figure 3.2). Menu options can be triggered by typing a phrase that contains "menu" anywhere within the phrase (top right). The user can click on the menu options or type a phrase in the "Ask something" text-input box. Below the list of menu options, the user is shown adding a blood pressure measurement and receiving a brief motivational message (middle left). Monthly graphs of blood pressure tracking can be viewed and shared with a provider (middle right). The user can view a list of medications, instructions, side effects (bottom left), and refill dates (bottom right).


Medicagent
Hi
Hi there! I'm Medicagent, your virtual medication assistant. I can provide you with information about your medications, help you track your blood pressure, schedule appointments, and more.
You can type "Menu" at any time to see a menu of options I can help you with. You can also use the chat below to ask questions.
Go ahead and type "Menu" to see what I can do!
Ask something...


Medicagent
You can also use the chat below to ask questions.
Go ahead and type "Menu" to see what I can do!
menu
Here are some options I can help you with.
Medications
Refills
Health and Blood Pressure Tracking
Appointments
Healthy Resources
Ask something...


Medicagent
Medications
Refills
Health and Blood Pressure Tracking
Appointments
Healthy Resources
Please add my blood pressure of 120/80 on 9/1/20
To confirm, would you like your blood pressure of 120/80 to be added on September 1, 2020?
Confirm BP
Great! Keep it up 👍
Ask something...


Medicagent
Here are your blood pressures for August 2020.
Blood Pressure - August 2020
160
140
120
100
80
60
40
20
0
Blood Pressure (mmHg)
Systolic BP
Diastolic BP
8/1
8/8
8/15
8/22
8/29
Date
Add blood pressure
Share blood pressure
Ask something...

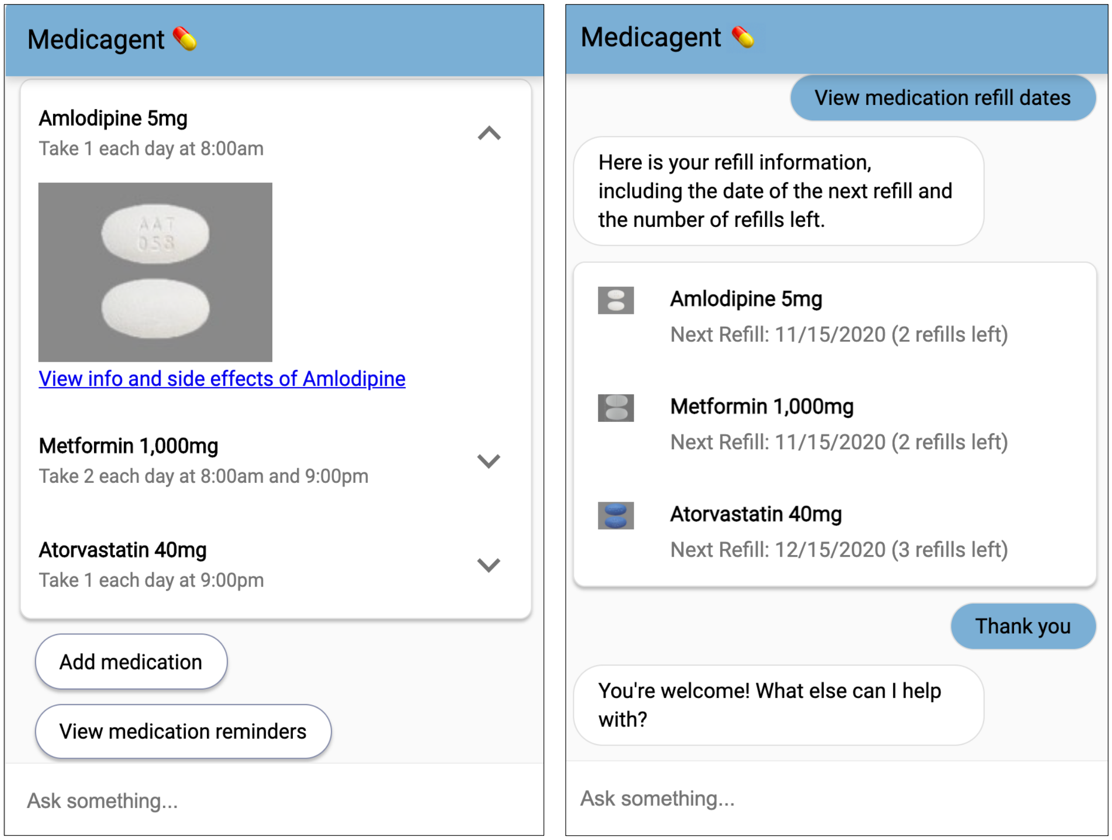


Figure 3.2 Screenshots of the user interface

Chat Engine and Database

Google's Dialogflow chat engine utilizes two algorithms to match natural language utterances to intents: rule-based grammar matching and machine learning (ML) matching[163]. This hybrid architecture runs both algorithms concurrently and chooses the one with the highest score. Dialogflow scores potential intent matches with an intent detection confidence score ranging from 0 (very uncertain) to 1 (very certain). We used a classification threshold of 0.3, which is the default setting in Dialogflow. If the highest scoring intent was 0.3 or above, an utterance was matched to the intent. If the score was less than the 0.3, a fallback intent was triggered which contained variations of "I'm sorry, I didn't understand that. Try again or type 'menu' to see the options I can help you with."

The chat engine was iteratively trained with example-based data of user utterances. We used approximately 10-15 training phrases for each intent as ML matching expands the list with additional similar phrases. The prototype underwent extensive pilot testing from health professionals and informatics researchers during development. We iteratively added utterances identified from the chat logs to the training phrases to improve system performance. Table 3.1 contains example training phrases and chatbot responses or parameter prompts for each intent. The actual training data contained additional variations of these phrases. Algorithms were updated when changes were made to utterances, intents, or parameters.

After an utterance was matched to an intent, parameters were extracted and data were used determine the necessary action, such as checking the database to determine if an appointment slot is available (Figure 3.3). Parameter prompts were used if the user did not fill all of the needed parameters from their initial utterance. These prompts contained examples of recommended input formats for dates, times, doses, and blood pressure values, and regular expressions were used to accept a variety of formats. As there was no pre-defined decision logic, users could input varying amounts of information ranging from one word that fulfilled no parameters to a question that fulfilled all parameters. For example, to set up a medication reminder, a user could input “set up medication reminder” and follow the parameter prompts to input the medication, date, and time. Alternatively, a user could input “set up a medication reminder for Lisinopril every day at 8:00am” and then confirm the medication reminder with the information included. If the initial utterance was matched to an intent but a parameter was not understood, a fallback message for each intent was used, such as “Sorry, I don’t recognize that medication. Would you like to try entering the medication again?" or "Hmm… It looks like you don't have any blood pressure recordings during that time. Try to enter another time." For this initial prototype, Firestore contained limited examples of appointment times, medications, refills, and blood pressure data.

Table 3.1 Example training phrases and parameters by intent

Intent	Training Phrases	Parameter Prompts or Response
Menu	Menu Menu options Help Go back What can you do?	Here are some menu options I can help you with. Click below on the topic you are interested in!
Medication list	Medication list View my meds What is on my med list? I would like to see my list of medications	Here are the medications you are currently taking.
Medication side effects	Side effects Learn more about Amlodipine What are the side effects of Hydrochlorothiazide?	To learn more about a medication, such as the side effects, please click on the medication below.
Add medication	Add med I would like to add a medication to my list Add Lisinopril 10mg per day at noon	Which dose of [medication] would you like to add? (e.g., 25 mg) What time would you like to take [medication]? (e.g., 11 am)
Medication reminder	Set up reminder Reminder schedule Could you give me a reminder for Lisinopril every Saturday and Sunday at 10 am?	How often would you like a reminder for [medication]? (e.g., everyday, weekends) What time would you like the reminder? (e.g., 9 am)

Refills remaining	Refills How many refills left? How many refills and the date of next refill for Lisinopril?	Here is your refill information for all of your medications, including the date of the next refill and the number of refills left.
Blood pressure values	View bp Can I see my blood pressure? Look at blood pressures for July 2020	Which month would you like to view your blood pressure? Here are the months that you have blood pressure data available.
Add blood pressure	Add bp Add my blood pressure My blood pressure was 125/75 Add 120/80 for 9/1/2020 Add 120/80 that was taken on Sept 1, 2020	What was your blood pressure? (e.g., 120/80) What date did you take your blood pressure? (e.g.,1/1/2020)
Share blood pressure	Send bp Share my blood pressure Send blood pressure measurements to (name of provider) Can you send my bp for Aug 2020 with Dr. Gupta?	Which month would you like to share your blood pressure? Here are the months you have blood pressure data available. Which provider would you like to share your blood pressure with?
Schedule appointment	Appointment I would like to schedule an appointment with Dr. Smith Schedule appointment 10/15/2020 at 3:00pm Make an appointment with Dr. Gupta on November 16 at 2pm	Which provider would you like to make an appointment with? The following appointment times are available with [provider]. Please select the time you are available or you can choose another date.
Recipe	Recipes Healthy food I would like to find a healthy dinner meal	Here are some healthy recipes which are updated each week.

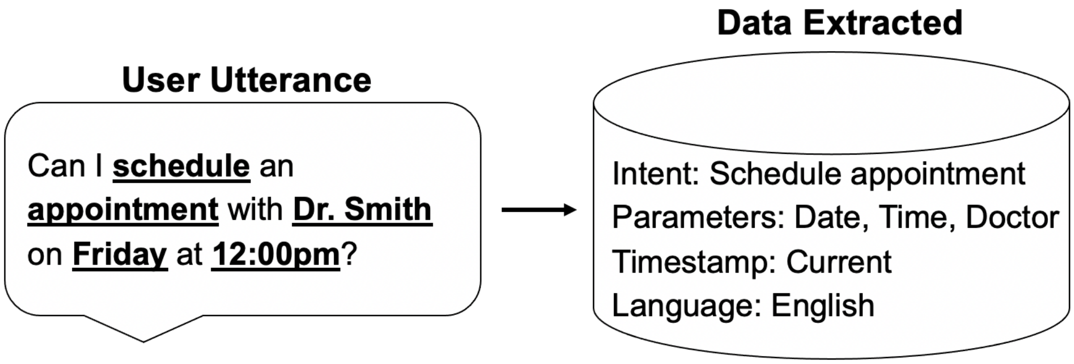


Figure 3.3 Data extraction in Dialogflow*
*When a user requests to schedule an appointment, Dialogflow matches the user's utterance to an intent. Parameters of the utterance are extracted to provide context to the user's intent. If a parameter is not filled from the utterance, the user is prompted to provide the information. Additional metadata is extracted such as the timestamp and language.

Study Design

We used standardized methods for usability studies of mobile health (mHealth) systems[164]. Effectiveness, efficiency, and satisfaction metrics in the International Organization for Standardization (ISO) 9241-11 standards were used to assess usability of the Medicagent prototype[63]. Usability testing was conducted virtually using one-on-one Zoom sessions due to the COVID-19 pandemic which restricted in-person studies. Effectiveness was measured by the percentage of tasks completed, user task error rate, and system error rate (i.e., unrecognized utterances). Efficiency of use was measured by the number of clicks, utterances, and duration of interaction. Satisfaction was measured by the System Usability Scale (SUS). Qualitative feedback was used to identify strengths and shortcomings of Medicagent and general feedback on the experience of using the chatbot. This study was reviewed and considered exempt by the University of North Carolina at Chapel Hill Office of Human Research Ethics Institutional Review Board.

Sample and Sampling

Eligible participants were adults 18 years and older with self-reported hypertension who took at least one hypertension medication. Participants had to have access to a computer, be able to use it without accessibility tools (e.g., did not require special software such as screen readers or alternative controls), and be willing to allow audio, video, and screen recordings of the Zoom session. Participants had the option to attend a brief Zoom set-up session before the day of the testing session to practice sharing their screen and communicating through the Zoom chat window, but we did not provide any specific tutorials or information about chatbots during these Zoom set-up sessions. 40% of participants attended this optional Zoom set-up session. Eligible participants were recruited using websites, e-mail list-servs, and flyers throughout hospitals, clinics, and community locations near Chapel Hill, NC. Recruitment materials contained a link to an electronic questionnaire to assess eligibility. Purposive sampling was used to select ten individuals[165] based on age, race, gender, education, and number of prescribed medications for representation across different sociodemographic and clinical characteristics. We aimed to have at least 50% of participants with the following characteristics: 65+ years old, minority race, male, education less than college, and taking at least three medications.

Procedures

There were four components of the testing session: 1) background questionnaire, 2) representative tasks within Medicagent, 3) usability questionnaire, and 4) brief semi-structured interview (see Appendix B for the full study protocol). One member (AG), who has been trained in usability testing and interviewing, conducted the usability sessions using a testing guide. The testing guide and questionnaires were pilot tested with members of the study team. There were no established interviewer-participant relationships, and the interviewer had no conflicts of interest with participants. None of the participants had participated in the prior studies.

First, participants completed the background questionnaire on topics of sociodemographics, medical history, experience with technology, health literacy[134], medication self-efficacy[135], and barriers to medication adherence[136]. Next, the task component comprised five hypothetical data entry and five data retrieval tasks based on hypertension and medication self-management processes. All tasks include synthetic information (i.e., medications, refills, blood pressure values) and did not contain participants' actual treatment regimens. Representative tasks included 1) viewing the medication list, 2) finding medication side effects, 3) adding a medication, 4) updating a medication reminder, 5) viewing refills, 6) viewing monthly blood pressure measurements, 7) adding a blood pressure measurement, 8) sending blood pressure to a health professional, 9) scheduling an appointment, and 10) finding a healthy recipe. While completing these tasks, participants were asked to describe their thoughts, feelings, and actions using concurrent think aloud methods to provide insights into their cognitive processes[166].

Participants were verbally introduced to Medicagent during the testing session by the interviewer as follows: "Medicagent is a virtual medication assistant focused on hypertension. The virtual assistant can provide you with information and help you keep track of your medications, refills, or blood pressure. It could also help you schedule appointments with your doctor, send you appointment reminders, or provide some health coaching. It could be accessed on your phone, tablet, or computer. The pictures and information that you see in Medicagent are for demonstrative purposes during our testing session today." No other information or tutorials were provided. Participants then accessed Medicagent from their personal computer on a website which contained a collapsed list of tasks on the left side and the chatbot hosted in the lower right side of the webpage. Participants completed an example task which initiated their interaction with Medicagent ("You would like to start chatting with Medicagent. Type Hi to get started."). The interviewer asked participants to click on a single task at a time, and all tasks were completed in the same order during each testing session. Participants did not receive help or prompting from the interviewer while completing tasks.

Following task completion, participants completed the SUS which is a reliable 10-item questionnaire used to measure the usability of a system[167]. Lastly, participants were asked about their perceptions, acceptance, and suggestions for modifications in a brief semi-structured interview using an interview guide. Each session lasted approximately 1.5 hours, and participants were provided with a $50 electronic gift card upon completion. The entire session was video and audio recorded over Zoom, and the interviewer took brief notes during and after the session. All ten participants completed the full testing session. There was no additional contact with participants.

Analysis

Questionnaire scores were calculated following standard scoring methodology from the validated instrument and summarized with descriptive statistics. For the 3-item Brief Health Literacy Screener, total scores range from 3 to 15 and any response greater than 3 for any question indicates inadequate health literacy[134]. The Patient-Reported Outcomes Measurement Information System (PROMIS) Self-efficacy for

Managing Medications and Treatments Short Form 8a (2016) was scored using the PROMIS HealthMeasures Scoring Service, where raw scores were converted into T-scores with a mean of 50 (SD=10) with higher scores representing greater self-efficacy[135]. The ASK-12 scores range from 12 to 60 with higher scores representing greater barriers to adherence[136]. The SUS score ranges from 0 to 100 with higher scores representing greater usability[167]. Prior approval was obtained for use of the 3-item Brief Health Literacy Screener by Chew et al (2004) and Adherence Starts with Knowledge 12 (ASK-12) by Matza et al (2009)[136]. Users' clicks, utterances, and task durations were extracted from Google Cloud Dialogflow analytics, and were verified through watching the Zoom sessions. User data were summarized using descriptive statistics. The session was transcribed verbatim using Zoom transcription, and the think-aloud and interview responses were analyzed by a trained qualitative researcher (AG) using a thematic analysis approach. The reviewer inductively identified and applied thematic content codes to identify the strengths and weaknesses of Medicagent across all participant narratives. Content codes were organized thematically to describe the major themes, subthemes, and illustrative quotes. Themes were discussed and revised with members of the research team.

Results

Participant Characteristics

Ten participants completed the usability testing session. The average age of participants was 60 years, 50% were female, 50% were Black, and 50% had at least a college education (Table 3.2). 80% had been diagnosed with hypertension at least five years ago, and participants were taking an average of four medications. Half had used a chatbot before and reported using Apple Siri, Amazon Alexa, Google Assistant, and customer service chatbots from websites. The majority had scores above the U.S. population average for medication self-efficacy (51) and felt that the greatest barriers to adherence were treatment beliefs (i.e., “I feel confident that each one of my medicines will help me”).

Table 3.2 Sample characteristics

Characteristics	**n (%)**
Age (mean=60, SD=10)	
35 – 44 years	1 (10)
45 – 54 years	2 (20)
55 – 64 years	2 (20)
65+ years	5 (50)
Gender	
Female	5 (50)
Male	5 (50)
Race	
Black or African American	5 (50)
White or Caucasian	4 (40)
Other	1 (10)
Ethnicity	
Not Latino/Latina	8 (80)
Latino/Latina	2 (20)
Education	
High school, GED, or less	2 (20)
Some college	3 (30)
College graduate or more	5 (50)
Household Income	
$35,000 – $49,999	1 (10)
$50,000 – $74,999	6 (60)
$75,000 or more	2 (20)
Did not report	1 (10)
Comorbidities, mean (SD)	2 (1)
Years with Hypertension	
1 – 2 years	2 (20)
5 or more years	8 (80)
Number of Prescription Medications, mean (SD)	4 (2)
Confidence Blood Pressure is Under Control	
Not confident at all	2 (20)
A little confident	1 (10)
Somewhat confident	2 (20)
Very confident	4 (40)
Completely confident	1 (10)

Internet Use Several times a day Almost constantly	 4 (40) 6 (60)
Device Use Smartphone Basic cell phone Tablet Computer	 9 (90) 1 (10) 5 (50) 10 (100)
Ever Used a Chatbot Yes No/Don't Know	 5 (50) 5 (50)
Health Literacy Level[134] Adequate	 10 (100)
Medication Self-Efficacy[135], mean (SE)	51.0 (4.1)
Barriers to Adherence[136], mean (SD) Treatment Beliefs Behaviors Inconvenience/Forgetfulness Total Score	 9.1 (3.2) 7.2 (2.2) 6.7 (2.5) 23.0 (6.2)

Summary of Tasks

Effectiveness

Nearly all tasks (98%) were successfully completed (Table 3.3). Two participants made errors that prevented task completion, which included inputting the incorrect medication and not confirming a new medication was added to the medication list. Among the ten participants, a total of 252 button clicks and 128 utterances were made. 8.6% (11/128) of utterances were not successfully mapped to an intent. These errors resulted from unrecognized spelling or formatting of dates, times, and blood pressure values. In these cases, Medicagent prompted the user to re-enter the information with a suggested format, and all users were then able to complete the corresponding task. Examples of participants utterances and the corresponding intents are shown in Table 3.4.

Table 3.3 Task summary (Mean, SD)

Task	Duration in Seconds	Button Clicks	Utterances	User Errors (error rate)[3]	Chatbot Errors (error rate)[4]
1. Find the list of current medications.[1]	46.1 (25.4)	1 (0)	0 (0)	0	0
2. Find 2 of the side effects of Amlodipine.[1]	69.1 (25.3)	3 (1)	1 (1)	0	0
3. Add 10mg of Lisinopril every day at noon to your medication list.[2]	131.6 (84.1)	2 (1)	3 (1)	1 (10%)[5]	4 (12.9%)[7]
4. Update your medication reminder schedule for Amlodipine to remind you to take it on weekends at 10am.[2]	120.7 (52.6)	4 (1)	2 (1)	1 (10%)[6]	1 (4.5%)[7]
5. View how many refills are left and the date of your next refill for Amlodipine.[1]	60.8 (36.1)	1 (1)	1 (1)	0	0
6. Find your blood pressure values for the month of August 2020.[1]	54.1 (20.8)	3 (3)	1 (1)	0	0
7. Add a blood pressure measurement of 120/80 that was taken on September 1, 2020.[2]	100.1 (66.8)	2 (1)	3 (1)	0	6 (23.1%)[7,8]
8. Share your blood pressure measurements for the month of August 2020 with Dr. Smith.[2]	69.3 (29.6)	4 (1)	1 (1)	0	0
9. Schedule an appointment with Dr. Smith for Tuesday, November 17th at 1pm. [2]	43.8 (24.1)	4 (2)	1 (1)	0	0
10. Find a healthy dinner recipe.[1]	60.2 (28.8)	2 (1)	1 (1)	0	0
Total (mean)	179 min (17.9 min)	252 (2.5)	128 (1.3)	2 (2.0%)	11 (8.6%)

[1]Data retrieval task,[2]Data entry task,[3]User error rate was calculated per task by: number of participants who did not successfully complete the task/total number of participants,[4]Chatbot error rate was calculated per task by: number of unrecognized utterances/total number of utterances, [5]Type of user error: did not confirm medication was added, [6]Type of user error: inputted incorrect medication, [7]Type of unrecognized error: date/time format, [8]Type of unrecognized error: blood pressure format

Table 3.4 Examples of participant utterances

Intent	Participant Utterances
Menu	*"Menu"* *"See menu"*
Medication list	*"Medications"*
Medication side effects	*"What are the side effects of Amlodipine?"*
Add medication	*"How do you add a medication? Thanks."* *"I want to add 10mg of Lisinopril once a day at 12 noon"*
Medication reminder	*"How do I get to the Amlodipine reminder for the weekend?"* *"Add reminder"* *"Update medication reminder for Amlodipine"*
Refills remaining	*"Refills left on amlodipine"* *"How many refills and when are the refills due?"* *"How many refills do I have left on Amlodipine and when is my next refill due?"*
Blood pressure values	*"Blood pressure values for the month of August 2020"* *"For the month of August what was my blood pressure values"*
Add blood pressure	*"Add my blood pressure measurement of 120-80"* *"Add 120/80 that was taken on September 1, 2020."* *"120/80 Sept"*
Share blood pressure	*"I wish to retrieve those August 2020 blood pressure measurements to send to Dr. Smith"* *"Send Dr. Smith blood pressure measurements"* *"Please share my blood pressures for the month of August with Dr. Smith"*
Schedule appointment	*"Schedule an appointment for me with Dr. Smith on Tuesday November 17th at 1pm"* *"Schedule an appointment with Dr. Smith for Tuesday, November 17 1pm"*
Recipe	*"Find a healthy dinner recipe"*

Efficiency

During testing sessions, participants spent an average of 18 minutes (SD=10) interacting with Medicagent (Table 3.3). Data retrieval tasks were completed faster on average (58 seconds) as compared to data entry tasks (93 seconds). Adding a new medication took the most time on average (132 seconds), and was the first data entry task. Scheduling an appointment took the least time on average (44 seconds), and it was the last data entry task. As users continued through the tasks, data entry task duration decreased, whereas data retrieval duration somewhat increased. This may be due to two data entry tasks that involved clicking on a hyperlink which opened up a new tab. Several participants found it difficult to navigate between tabs on their computer. Overall, data entry tasks had more button clicks and utterances.

Satisfaction, Strengths, and Shortcomings

Medicagent achieved a mean SUS score of 78.8 out of 100 (Table 3.5). Scores below 50 are generally considered not acceptable, 50 to 70 as marginal, and above 70 as acceptable[168, 169]. Participants reported a number of strengths and shortcomings of Medicagent while thinking aloud and during interviews (Table 3.6). Overall, most participants reported Medicagent was easy to use and enjoyable. Several felt it became simpler to use the chatbot as the tasks proceeded. A few participants also found it useful to confirm adding a medication reminder or blood pressure measurement. Nearly all had positive attitudes towards the visuals, such as images of medications and blood pressure charts.

For the shortcomings, several desired a menu button to aid in navigation. Many felt typing ‘menu’ (or any utterance including the word ‘menu’) was not a natural interaction. When probed about the type of desired menu, participants described a hamburger-like option that included dropdown options of common tasks. Additionally, many felt a back button would be useful in case an error was made which would allow them to go back to the previous step. As two of the tasks involved clicking on a hyperlink that opened up a new tab, some preferred to complete all self-management tasks within the chatbot. Several thought the chatbot could be improved by adding a persona of a health professional, such as a medical avatar. A few also desired to interact with Medicagent through voice or a combination or voice and text.

Table 3.5 System Usability Scale scores*

Questionnaire Items	**Mean (SD)**
1. I think that I would like to use this system frequently.	4.0 (0.8)
2. I found the system was unnecessarily complex.	2.0 (0.8)
3. I thought the system was easy to use.	4.0 (0.7)
4. I think that I would need the support of a technical person to be able to use this system.	1.5 (0.7)
5. I found the various functions in the system were well integrated.	4.1 (0.9)
6. I thought there was too much inconsistency in this system.	1.7 (0.8)
7. I imagine that most people would learn to use this system very quickly.	4.1 (0.9)
8. I found the system very cumbersome to use.	1.7 (0.8)
9. I felt very confident using the system.	4.0 (1.1)
10. I needed to learn a lot of things before I could get going with this system.	1.8 (0.8)
Total SUS Score	78.8 (15.9)

*1=strongly disagree and 5=strongly agree

Table 3.6 Pervasive themes for usability strengths and shortcomings of Medicagent

Strengths	**Representative Quotes**
Easy to use	*"There was a short learning curve, and it was easier as the tasks went on as I remembered what the menu items were."*
Enjoyable	*"It even gives me a little motivation to continue to exercise and take my Lisinopril…this is quite fun."*
Visuals	*"The graph is very interesting with the systolic and diastolic lines."* *"I can see the pills. That's very useful. There are certain pills that I have that look exactly alike which can be a problem."*
Shortcomings	**Representative Quotes**
Menu button	*"Typing the word 'menu' is not something I have used in any other app before…maybe have those three lines that are usually used for menus in the top right or left corner."* *"I was scrolling up and down a lot to find things. It was like looking for an old text message…It would be a good idea to have a home button."*
Back button	*"It seems like a back button is needed. I made an error and had to go all the way through the questions again."*
Navigating between tabs	*"Going to so many different screens sometimes can confuse me and then getting back to the chatbot screen was not that easy."*
Persona	*"I was hoping it would be an animated character instead of just words…He could wear a medical hat or white coat."* *"I was expecting something that looks like a doctor that talks…Voice interaction would be really cool."*

Discussion

Principal Findings

We designed, developed, and evaluated the usability of a hypertension self-management chatbot prototype, Medicagent. Almost all (98%) tasks were successfully completed during testing. The time spent completing tasks decreased with each additional data entry task, but varied with the data retrieval tasks that included hyperlinks to two websites. This suggests that participants may have become more familiar with additional experience with the chatbot as tasks progressed, which can be expected. Similarly, some reported during interviews that there was a short learning curve for using the chatbot, and Medicagent received an average SUS score of 78.8 which demonstrates acceptable usability. In comparison to other text-based health chatbots, mean SUS scores have been 81.8[71] and 88.2[170]. These studies evaluated usability of mental health chatbots, and the mean age of participants was younger than in our study. To date, most usability evaluations of chatbots have not used the SUS or other validated instruments so there is limited comparison data. Several participants had difficulties navigating without standard features like menu and back buttons that are usually found in websites, and many reported the desire for a health professional identity and personality to be embodied in the chatbot. Through our usability testing and interviews, we identified three main components that may facilitate usability of chatbots for self-management: interaction flexibility, graceful degradation (i.e., the ability of the system to tolerate failures)[171], and a medical professional persona.

Interaction Flexibility

In our study, we observed two primary types of interaction styles driven by buttons or utterances. Participants who mainly used buttons began completing tasks by clicking on a menu button, while those who used utterances completed tasks using the text-input box. If all parameters were not filled on the initial utterance, Medicagent prompted the user for the missing parameters. For example, one participant typed *"Add my blood pressure measurement of 120-80"* and was prompted for the date the blood pressure was taken. A few participants completed the first few tasks using the button approach and then switched to typed utterances as the tasks progressed. Prior studies have found that users prefer to have the option to input free-text or use buttons[172]. This suggests that both input mechanisms should be available to afford flexibility to end-users. However, unrecognized free-text responses may pose potential patient safety risks if appropriate safety measures are not in place[44, 88]. Confirming user input may help mitigate some of these concerns, and several participants found the confirmations to validate their inputs to be helpful.

For those who preferred navigating from the menu, a menu icon at the top was desired to select from a dropdown list of options. Several participants felt that a back button would be useful to navigate back to a point immediately prior in the conversation instead of re-typing or re-querying. Some participants also wanted the flexibility to communicate through both voice and text. This may be due to half of participants in this study reporting previous use of voice assistants such as Apple Siri, Google Assistant, or Amazon Alexa. This suggests that users have diverse interaction preferences for self-management chatbots, which may be addressed through various visual cues (i.e., menu icons, buttons, text-input boxes) and multimodal interfaces.

Graceful Degradation

Participants inputted a variety of formats for dates and blood pressure values, and Medicagent re-promoted them when a misspelling or formatting issue prevented the parameter from being filled. For completely unrecognized utterances, Medicagent responded with variations of "*I'm sorry, I didn't understand that. Try again or type 'menu' to see the options I can help you with.*" However, some participants felt these error messages were too generic. Handling unrecognized errors gracefully is particularly important for chatbots because people generally perceive robots as intelligent and competent[173, 174], and may have less tolerance for such errors. Erroneous agents are perceived as less reliable which negatively affects task performance[173, 174]. At this early stage with limited health care corpora for training data, it is unlikely a chatbot would be able to recognize the vast number of possible free-text inputs for all situations. Thus, adding context to error messages to enable users to better understand the cause of the error may reduce user frustration[175]. Incorporating features that may ease data entry, such as a calendar of dates to select, could be helpful to minimize unrecognized utterances arising from data entry errors[155].

Medical Professional Persona

Several participants wanted the chatbot to have an avatar with visual characteristics of a health professional, such as a white coat or medical hat. A few felt these attributes would help to establish credibility within a system that tracked medications and provided virtual health coaching. Embodied conversational agents (ECAs) are computer-based characters that emulate face-to-face conversations by using speech and nonverbal characteristics, such as facial expression and hand gestures[176]. Prior studies of ECAs have simulated a health provider for diet or exercise coaching and review of hospital discharge materials[46, 53, 177, 178]. Overall, these have been received positively by patients and increased adherence to treatment regimens[46, 53, 177, 178]. It is possible that some participants were used to interacting with chatbots by their names, such as “Hey Siri” or “Hey Alexa” and may have wanted similar personifications in Medicagent. Consequently, including additional visual attributes, such as those in ECAs, may be beneficial in establishing patient rapport and improving satisfaction of the chatbot, which could lead to better engagement.

Implications for Health Care and Research

This research has several implications for the field of chatbots in health care and research. First, we provide an overview of our design process and highlight several components that may enrich the user experience of health-related chatbots which may provide some insights for chatbot development in the context of other chronic conditions requiring self-management. We also describe the development architecture, which could be used as a framework in future research to develop chatbots on Google Cloud. We used a labor-intensive process of iteratively training the chat engine through extensive pilot testing, followed by assessing the chat logs for unrecognized utterances or incorrect responses, and then adding utterances and intents to the training data. Similar manual processes have been reported in other chatbot development studies to fine-tune the system[179]. Zand et al (2020) used natural language processing to develop a chatbot knowledge base by categorizing electronic messaging data between patients and providers[180]. Patients' questions and concerns were mapped to categories, such as symptoms, medications, appointments, and laboratory investigations[180]. Similar approaches may be useful to inform the knowledge base for self-management conversational agents. Patient and provider validated and open-source example-based training sets could be used in the initial development phase then tailored appropriately for cultural and linguistic differences across patient populations. While there have been a number of corpora analyses on patient and provider communication[181-183], most corpora are not publicly available due to confidentiality issues. During the COVID-19 pandemic, several commercial services, such as Microsoft Azure and Amazon Web Service, released built-in frameworks for symptom checkers and other medical content[184, 185] that may also be useful an initial starting point in the design and development of future chatbots.

For our prototype, we used hypothetical data and did not include protected health information (PHI), which would require a Google Cloud Business Associate Agreement. These agreements ensure Health Insurance Portability and Accountability Act (HIPAA) compliance, and Dialogflow services are reported to be covered under these agreements[92]. As conversational interfaces are becoming more widespread messaging platforms and connecting to various APIs, it is vital to ensure confidentiality and privacy of PHI. Lastly, many health care chatbot studies do not use validated instruments which limits the ability to generalize findings. However, our usability evaluation provides data to compare performance in chatbot research.

Limitations

Our study sample was limited to individuals with a computer from a single geographic region in the Southeast, and all participants had adequate health literacy and high levels of medication self-efficacy. Prior research suggests that our sample size is adequate for usability testing[165], but we may not be able to draw definitive conclusions about the strengths and shortcomings of chatbots to support hypertension self-management. While concurrent think aloud methods are valuable for real-time feedback, task time varied based on the amount of feedback provided by participants. So, these task durations may not be representative of actual task times. All tasks were completed in the same order, which may present ordering bias in which users perform worse on the initial tasks while adjusting to the testing environment and being observed. A couple of participants also expressed frustration with pronunciation of the medications while thinking aloud (i.e., Lisinopril and Amlodipine), which could have decreased overall satisfaction during the testing session. Due to the COVID-19 pandemic, the usability testing was conducted remotely over Zoom on participants' personal computers. Although the website that contained Medicagent was consistent across all participants, there were variances in Internet speed, amount of personal notifications received, and users may have had a mouse or used a track pad. However, one benefit of conducting remote usability testing was that it allowed us to observe interactions within a user's natural environment where self-management tasks and interactions with the chatbot would typically take place.

Conclusion

In the emergent field of health care chatbots, we describe the design, development, and usability evaluation of one of the first known chatbots focused on hypertension and medication self-management, which was found to have high user acceptance and good usability. Flexibility of various interaction styles, handling unrecognized utterances gracefully with contextual error messages, and having a credible health professional persona to increase engagement were highlighted as design features that could facilitate usability and navigation within chatbots for self-management. This research contributes towards a better understanding of how patients with hypertension interact with a self-management chatbot, which may help to inform how designers could improve patient experience and promote engagement in self-management tasks in future research. Additional usability research for chatbots should investigate the appropriateness of chatbot responses within the context of self-management and how users' interactions could optimize their self-management goals.

CHAPTER 4: CLUSTERING PHYSICAL ACTIVITY PATTERNS FROM WEARABLE DEVICES

Background and Significance

Inflammatory bowel diseases (IBDs), which are comprised of Crohn's Disease (CD) and ulcerative colitis (UC), are chronic intestinal disorders of the gastrointestinal tract. IBDs are characterized by cycles of active and dormant states of inflammatory immune response with symptoms such as abdominal pain, diarrhea, or fatigue[186]. Physical symptoms are frequently accompanied by stress, anxiety, depression, or diminished quality of life and may be exacerbated by common immunosuppressive therapies or corticosteroid treatments[187]. Allowing patients with IBDs to directly report their symptoms or functional status provides an important clinical endpoint to understand the burden of disease[188]. Patient-reported outcome (PRO) measures of health status and functioning have been shown to correlate with IBD disease activity indices[189], which are used to assess clinically whether patients are in remission.

The association between disease activity and physical activity in patients with IBDs[186] is not well understood but positive impacts of exercise have been demonstrated across a number of other chronic conditions[190]. Evidence suggests that most patients reduce their physical activity following diagnosis of IBD[191]. Regular physical activity has contributed towards improvements in symptoms, although there may be differences in disease activity related to the duration, frequency, or type of activity (e.g., resistance, endurance, etc.)[186]. For example, moderate physical activity interventions, including cardiovascular training, strength training, and yoga, have shown generally positive effects on symptoms and inflammation[191-195]. However, high intensity or extended durations of exercise can lead to inflammation and exacerbate gastrointestinal symptoms[192, 196]. Examining these dimensions of physical activity relative to IBD disease activity and psychosocial outcomes has not yet been studied.

Estimates indicate that up to 45% of the U.S. population owns a wrist-worn wearable device (wearable)[197, 198], which objectively measure activity and allow monitoring of exercise over time. Physical activity phenotyping from these emerging devices may be useful to distinguish multifaceted activity patterns. Prior studies have derived activity groups based on various exercise dimensions, including ‘weekend warriors,’ ‘busy bees,’ ‘drivers,’ ‘insufficiently active,’ ‘cardio active,’ or ‘endurance athletes’[199-202]. These studies demonstrated the ability for clustering approaches to define groups with related physical activity patterns. Being able to identify cohesive physical activity phenotypes from multidimensional data is valuable as the current physical activity guidelines utilize multiple exercise dimensions, including aerobic activity intensity, muscle-strengthening activity, and duration[203]. There may be additional characteristics useful for assessing physical activity beyond these guidelines, such as steps or calories, which are often tracked by wearables. Few studies have used data from wearables to assess physical activity in patients with IBDs, though early findings suggest an association between physical activity and biomarkers for inflammation and disease activity[204, 205]. A better understanding of physical activity patterns and intensities could reveal unique lifestyle characteristics and insights into the disease’s impact on health status and functioning. These phenotypes could provide a more holistic view of a patient's physical activity or inform lifestyle or treatment interventions.

Objectives

We sought to cluster wearable device data to identify physical activity phenotypes and better understand the relationship between physical activity clusters and patient-reported outcomes for disease activity and psychosocial domains over time in patients with inflammatory bowel diseases.

Methods

Study Setting

The Crohn's and Colitis Foundation of America (CCFA) Partners study is an Internet-based cohort of adults (18+) with self-reported IBDs[206]. Participants in the cohort have access to an online portal where they can sync smartphone applications or wearables and complete biannual PROs questionnaires[24]. Wearables data are contributed through a bring-your-own-device model, where participants can sync any brand or device used to monitor physical activity (e.g., Fitbit, Garmin, Under Armour, etc.).

Data comes from the Precision VISSTA study, which uses CCFA Partners data to develop precision health recommendations for lifestyle behaviors to improve health outcomes. Participants were included in this analysis if they completed at least one questionnaire and contributed wearable device data from July 2011 to February 2020. This study was reviewed and considered exempt by the Institutional Review Board at the University of North Carolina at Chapel Hill.

Participants completed biannual questionnaires containing sociodemographic information and Patient-Reported Outcomes Measurement Information System

(PROMIS) short forms in the psychosocial domains (depression, anxiety, pain interference, sleep disturbance, social relationships, fatigue). Participants also completed disease activity questionnaires. CD disease activity was measured with the short Crohn's Disease Activity Index (SCDAI), which assesses severity of symptoms (i.e., stool frequency, abdominal pain, well-being) within the past 7 days[207]. SCDAI scores range from 0 to over 600 with scores <150 indicating clinical remission and ≥150 indicating active disease (mild activity: 150-219, moderate activity: 220-450, severe activity: >450)[207]. UC or indeterminate colitis (IC) disease activity was measured using the Simple Clinical Colitis Activity Index (SCCAI) to measures the severity of symptoms (i.e., frequency of bowel movement during the day and night, urgency of defecation, blood in stool, well-being, extracolonic features) within the past 7 days[208]. SCCAI total scores range from 0 to 19. Scores <2.5 correlate with remission, whereas scores ≥2.5 correlates with active disease[209]. PROMIS questionnaires were scored using standardized T-score distributions where a score of 50 (SD=10) represents the population mean of the reference population. T-scores represent the concept being measured (e.g., more fatigue, more social relationships, etc.).

Data Preprocessing

Physical activity metrics varied across device brands (e.g., Fitbit, Garmin, Under Armour, etc.), so matrices were created to illustrate the capabilities of each device. Measurement systems also differed, and features were standardized into minutes for duration and miles for distance. Tukey's method was used to assess outliers, which were considered observations 1.5 times less or greater than the lower and upper quartile ranges, respectively[210]. Unrealistic values were removed (i.e., activity duration of 24 hours).

To determine the period of physical activity prior to completing a questionnaire to include in our analysis, Spearman's correlation coefficients were calculated to measure the strength of the relationship between physical activity (steps) and disease activity (SCCAI and SCDAI). Coefficients were calculated at various weeks prior to completing a questionnaire (i.e., weeks 1, 2, 4, 6, 8, 12, and 24) with at least 50% of days with data returned during the time period. The period with the highest correlation between steps and disease activity was selected, which was 6 weeks. During this 6-week period, the correlation coefficient for SCDAI was -0.19 ($p<0.01$) and -0.14 ($p<0.01$) for SCCAI. Thus, participants were included in the analysis if they had at least 21 days of activity data (non-consecutive) within the 42 days (6 weeks) prior to completing a questionnaire. Approximately 20% of participants (113/543) were excluded who had completed at least one questionnaire but did not have at least 21 days of activity data within the 6 weeks prior to completing the questionnaire. If participants completed more than one questionnaire and had at least 50% of activity data within the 6 weeks prior to the questionnaire, there could be multiple observations per participant. This allowed us to examine the change in consecutive questionnaire scores as participants moved to/from clusters.

Features included in the analysis were average ratios of the following: number of steps, moderate to vigorous activity duration (minutes), distance of activity (miles), number of calories burned, number of days device was used during weekdays, and number of days device was used during weekends. For example, the average number of steps was calculated using a ratio of the total number of daily steps taken and the number of days the participant contributed steps data. For each participant, features were averaged during the 6-week time period prior to completing each PROs questionnaire. K-means algorithm was iteratively run using combinations of these features, and the quality of clusters were evaluated using silhouette coefficients. The combination of features that produced the highest average silhouette coefficient was selected. This included three features: number of steps, minutes of moderate to vigorous activity duration, and distance of activity in miles. These features were moderate to highly correlated as demonstrated by Spearman's correlation coefficients (ranging between 0.63 and 0.95). We initially performed featuring scaling (z-score standardization and min-max normalization), which had little effect on the quality of clustering. Thus, the raw preprocessed data were used in the clustering.

Cluster Identification and Evaluation

K-means cluster analysis or Lloyd's method[211] was conducted to generate the physical activity groupings. K-means uses an iterative approach to assign each data point (x) to the closest centroid (c_i) within each cluster (C_i) by minimizing the average sum of squared Euclidean distance[212]. The equation below represents the K-means algorithm in which the sum of squared errors (SSE) or cluster scatter is minimized. K-means analyses were conducted using the Scikit-learn package in Python 3.7[213]. The following parameters were used: number of clusters=3; initialization=k-means++; number of initializations=10; maximum iterations=300; tolerance=0.0001; random state=none. The algorithm was initialized 10 times with the centroids initially selected randomly (random state=none), and the initial centroid distances were optimized with k-means++. K-means++ places centroids to be generally distant from each other, which has been shown to improve the speed of convergence[214]. When the difference in the centroids across two consecutive iterations was lower than the tolerance level (0.0001), the algorithm converged and iteration stopped.

$$\mathrm{SSE} = \sum_{i=1}^{K} \sum_{x \in C_i} dist\,(c_i, x)^2$$

where:

- K is the number of clusters
- x is a data point that belongs to cluster C_i
- C_i is the i^{th} cluster
- $dist$ is the standard Euclidean distance between two data points
- c_i is the mean (centroid) of cluster C_i

The number of clusters (K) was determined by calculating the SSE for different values of K. The value of K was selected where the change in SSE decreased and was becoming plateau, indicating additional clusters produced little value (“elbow point”). We iteratively evaluated the silhouette coefficients of two to five clusters and selected the model with the highest average silhouette coefficient, which was three clusters (Figure 4.1). The silhouette coefficient assessed the quality of clusters by measuring the amount of cohesion and separation within and among clusters. Ranging from -1 to 1, higher coefficients indicate better defined clusters[212]. Averages of silhouette coefficients were calculated for each cluster and for the average of each of the three clusters. Once the clusters were identified, data within each cluster were assessed to determine a label that appropriately represented each cluster’s attributes. Labeling clusters is a subjective process, and it is suggested to the name them in a way that is interpretable by the target audience[215]. We labeled the clusters as low physical activity, moderate physical activity, and high physical activity.

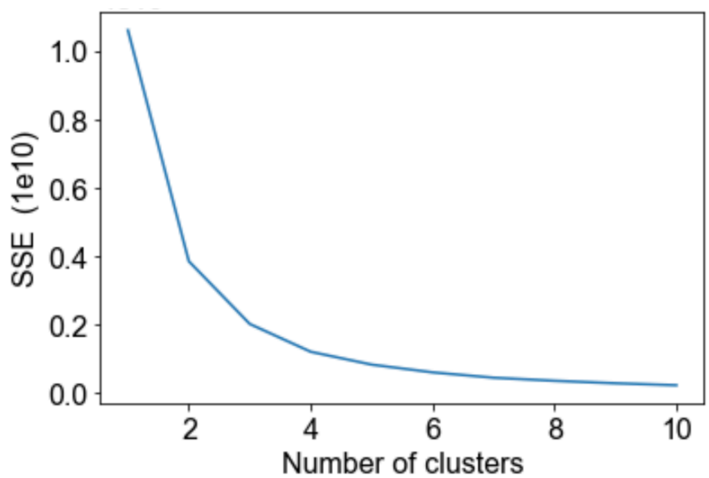


Figure 4.1 Sum of squared errors plot

$$s_x = \frac{b_x - a_x}{\max(a_x, b_x)}$$

where:

- a_x is the average distance from an individual data point x to all data points within the same cluster
- b_x is the average distance from an individual data point x to all data points in the nearest cluster

Association between Physical Activity and Health Status

Based on these clusters, we conducted a cross-sectional analysis to assess sociodemographic differences between the clusters and to compare means of disease activity and PROMIS psychosocial scores across clusters using one-way analysis of variance (ANOVA) tests. Participants were included at multiple time periods if they had multiple 6-week physical activity periods and completed multiple questionnaires. Therefore, the same participant could be in more than one cluster.

Lastly, longitudinal analysis was conducted on a subset of participants from the clusters who completed at least two consecutive questionnaires which span approximately a one-year time frame. Participants were grouped into categories related to staying in the same activity cluster or moving into another cluster across consecutive questionnaire time periods. For each move between clusters, mean differences in disease activity scores were compared using paired sample t-tests. For all analyses, a p-value < 0.05 was considered statistically significant.

Results

Participant Characteristics

The final analytic sample had 430 participants, of which 285 (66.3%) had Crohn’s Disease and 145 (33.7%) had ulcerative or indeterminate colitis (Table 4.1). Participants were primarily female, White, and attained at least a college degree. On average, age was 42.1 years, BMI was 25.8 (overweight), and duration of disease was 15.1 years. Participants in the sample completed an average of 3 questionnaires from 2015–2020. During the 6-week period prior to completing questionnaires, participants used their wearable device for an average of 37.3 days (88.9% of the 6-week period). This included approximately 89.7% (26.9/30) of weekdays and 86.7% (10.4/12) of weekends. The majority used a Fitbit device (86.3%). On average, participants took 7,893 daily steps, performed moderate to vigorous activity (i.e., active duration) for 41 minutes, traveled 3.5 miles, and burned 521 calories during exercise.

Table 4.1 Sample characteristics[1,2]

Characteristics, n (%)	**n=430**
Age, mean (SD)	42.1 (13.6)
Gender	
Female	318 (74.0)
Male	112 (26.0)

Race[2] White Black or African American Asian Other	 394 (95.2) 7 (1.7) 5 (1.2) 8 (1.9)
Ethnicity[2] Not Hispanic or Latino Hispanic or Latino	 405 (97.1) 13 (2.9)
Education[2] High School or Less Some College College Degree or More	 21 (5.0) 72 (17.1) 327 (77.9)
BMI, mean (SD)[2]	25.8 (4.8)
Smoking Status Ever Never	 121 (28.1) 309 (71.9)
Type of IBD Crohn's Disease Ulcerative or Indeterminate Colitis	 285 (66.3) 145 (33.7)
Duration of Disease (years), mean (SD)	15.1 (11.8)
Device Brand Fitbit Garmin Jawbone Under Armour	 371 (86.3) 43 (10.0) 12 (2.8) 4 (0.9)
Questionnaires Completed, mean (SD)	3.0 (2.0)
Days Device Used (maximum of 42 days), mean (SD)	37.3 (6.2)
Days Device Used on Weekdays (maximum of ~30 days), mean (SD)	26.9 (4.4)
Days Device Used on Weekends (maximum of ~12 days), mean (SD)	10.4 (2.1)
Daily Steps, mean (SD)	7,892.9 (2,752.7)
Daily Active Duration in minutes, mean (SD)	40.7 (38.3)
Daily Distance in miles, mean (SD)	3.5 (1.3)
Daily Activity-Related Calories, mean (SD)	520.7 (170.0)

[1]Demographic and clinical characteristics are from the baseline questionnaire;[2]Missing data: race=16; ethnicity=12; education=10; BMI=14 participants

Clusters of Physical Activity Patterns

K-means cluster analysis identified three cluster groups of physical activity that we labeled as low, moderate, and high activity (Figure 4.2). Among 430 participants, there were 1,255 total 6-week time periods. 423 (33.7%) periods were classified as low activity, 577 (46.0%) as moderate activity, and 255 (20.3%) as high activity. The three features used to generate the clusters and characteristics of each cluster are shown in Table 4.2. Overall, activity clusters were moderately defined (average silhouette coefficient=0.54). The quality of the clustering was highest in the low activity cluster (silhouette coefficient=0.60), suggesting that participants had the most similarities for steps, distance, and moderate to vigorous active duration within this cluster. Those in the high activity cluster had the most variance in their levels of exercise (i.e., steps, distance, and moderate to vigorous active duration) and least number of participants. Sociodemographic characteristics varied across clusters, and those in the low activity cluster were older, had higher BMIs, and longer disease duration when compared to the other clusters ($p<0.05$).

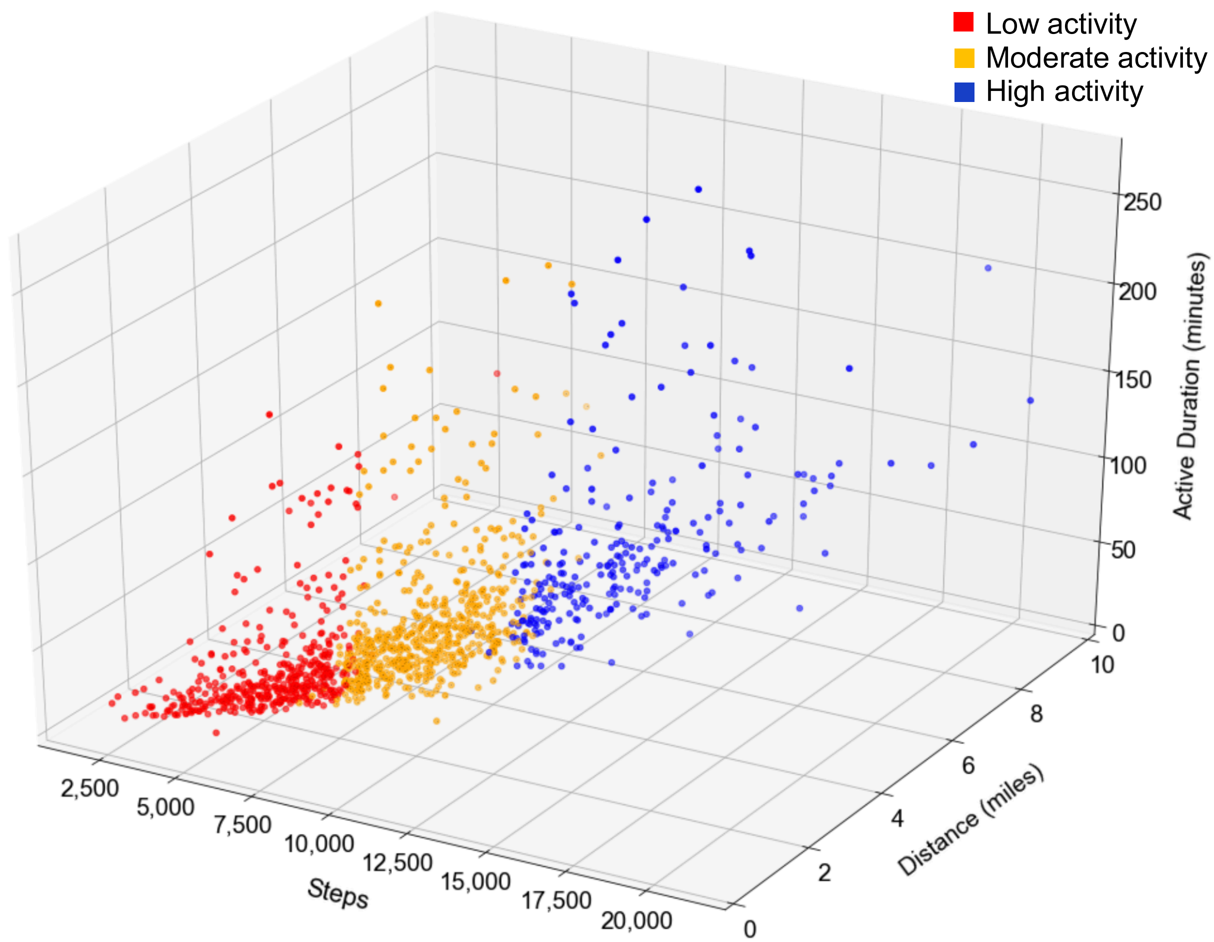


Figure 4.2 Physical activity clusters

Table 4.2 Cluster profiles

Characteristics	**Low Activity n=423**	**Moderate Activity n=577**	**High Activity n=255**
Evaluation			
Silhouette Coefficient	0.60	0.52	0.48
Features, mean (SD)			
Steps	5,000.3 (1,062.4)**	8,229.5 (1,033.0)**	12,319.4 (1,899.6)**
Distance (miles)	2.2 (0.6)**	3.7 (0.7)**	5.5 (1.1)**
Active Duration (minutes)	21.3 (27.0)**	40.6 (34.6)**	73.0 (46.4)**

Sociodemographics, mean (SD)			
Age	45.8 (13.8)**	43.5 (13.7)**	42.1 (12.6)**
BMI	27.8 (3.5)**	25.6 (4.5)**	23.6 (3.5)**
Duration of Disease (years)	17.4 (12.3)*	16.9 (12.0)*	14.8 (10.0)*

*p<0.05; **p<0.01

Association between Physical Activity and Patient-Reported Health Status

Across all disease activity scores and PROMIS psychosocial domains, patients in the low activity cluster had the worst scores and those in the high activity cluster had the best scores (Table 4.3). Scores varied significantly across clusters on the level of depression, pain, fatigue, sleep disturbance, social relationships, and short Crohn's Disease activity index (p<0.05). The largest differences between the low and high activity clusters for the PROMIS domains were scores in the domains of social relationships (5.0), fatigue (4.2), and pain (4.0). Although SCCAI scores did not reach significance for patients with UC or IC, those in the low and moderate activity clusters likely had some active disease activity during the 6-week period as scores ≥2.5 have previously demonstrated correlations with active disease[209]. Mean SCDAI scores were in remission (<150) in all clusters, but scores were highest in the low activity cluster[207].

Table 4.3 Patient-reported outcome scores across clusters

PRO Scores, mean (SD)	Low Activity n=423	Moderate Activity n=577	High Activity n=255
SCDAI[1]	133.3 (80.3)**	117.0 (71.9)**	102.2 (59.7)**
SCCAI[1]	2.9 (2.1)	2.6 (1.8)	2.3 (2.5)
Anxiety	50.4 (9.2)	49.3 (8.9)	49.2 (8.9)

Depression	48.6 (8.3)**	47.1 (7.8)**	46.9 (7.3)**
Pain	50.7 (9.5)**	48.5 (8.4)**	46.7 (7.6)**
Fatigue	54.9 (10.7)**	51.5 (10.9)**	50.7 (9.6)**
Sleep Disturbance	51.2 (7.5)**	49.4 (7.7)**	49.2 (7.8)**
Social Relationships	50.3 (9.7)**	54.0 (9.1)**	55.3 (9.1)**

[1]Missing data: SCDAI=44; SCCAI=78 questionnaires; *p<0.05; **p<0.01

Among the 246 participants who completed at least two consecutive questionnaires (726 total questionnaires), 67.8% (492/726) did not change physical activity clusters during 6-month periods (Table 4.4). As expected, there were no significant changes in mean disease activity scores among those who did not cross between clusters. For those who transitioned into another cluster, 15.8% (115/726) moved between (to and from) moderate and high activity, 15.2% (110/726) moved between low and moderate activity, and only 1.2% (9/726) moved between low and high activity clusters. Proportions of cluster movement were similar for patients with CD and UC/IC. There were significant associations between cluster movement and mean disease activity score for the three types of transitions. When patients with UC or IC transitioned from low to moderate activity clusters, disease scores decreased ($p<0.05$). For CD, when patients moved from moderate to high activity or high to moderate activity clusters, disease scores decreased or increased, respectively ($p<0.05$).

Table 4.4 Movement across clusters for consecutive disease activity scores

IBD Subgroup	Change in Physical Activity	Physical Activity Cluster Movement	n	Mean (SD) Change in Disease Activity
Ulcerative or Indeterminate Colitis (SCCAI)	Improved	Low → Moderate	21	-1.4 (2.4)*
		Moderate → High	19	-0.4 (1.7)
		Low → High	0	--

	Reduced	Moderate → Low	25	0.1 (1.2)
		High → Moderate	17	0.1 (2.1)
		High → Low	4	3.2 (3.8)
	No change	Low	51	0.1 (1.8)
		Moderate	76	-0.1 (1.3)
		High	38	-0.4 (1.8)
Crohn's Disease (SCDAI)	Improved	Low → Moderate	26	-0.5 (82.8)
		Moderate → High	38	-25.3 (64.5)*
		Low → High	1	0
	Reduced	Moderate → Low	38	7.6 (63.6)
		High → Moderate	41	20.6 (55.5)*
		High → Low	4	78.8 (111.8)
	No change	Low	124	0.5 (67.9)
		Moderate	140	-6.3 (59.5)
		High	63	8.2 (69.8)

[1]Data are based on 726 consecutive questionnaire pairs from 246 participants.; *p<0.05; **p<0.01

Discussion

Principal Findings

Our findings suggest physical activity phenotypes can be generated from consumer-based wearable devices in patients with IBDs. Most participants (46.0%) were clustered in moderate activity, 33.7% as low activity, and 20.3% as high activity. Sociodemographic characteristics varied across clusters, and those with low activity were older, had higher BMIs, and longer disease durations ($p<0.05$). We demonstrate positive correlations between physical activity and health status (i.e., IBD disease activity and psychosocial domains) in accordance with existing research[205, 216]. Patients in the low activity cluster had the worst scores across all PROs, and scores varied significantly on levels of depression, pain, fatigue, sleep disturbance, social relationships, and Crohn's Disease activity across clusters ($p<0.05$). Ulcerative colitis activity indices did not vary significantly across clusters, which may be due to patients in the low (mean SCCAI=2.9) and moderate (mean SCCAI=2.6) clusters who had some degree of active disease. Participants with low physical activity had the most homogeneity in exercise attributes (i.e., steps, distance, active duration), which suggests worsened health status might hinder variations in exercise.

When we longitudinally assessed changes in physical activity and disease activity scores across 6-month periods, exercise patterns mostly did not fluctuate. Approximately 68% of patients remained in their original cluster, and only 1% of patients transitioned to or from the furthest clusters of low and high activity. This indicates that exercise levels may not vary to extremes over time. As expected, mean disease activity scores among patients who remained in the same cluster did not change over time. However, disease activity scores varied among IBD subgroups for some patients who moved into different clusters. When patients with UC or IC transitioned from low to moderate activity clusters, disease scores decreased ($p<0.05$). When patients with CD transitioned from moderate to high activity or high to moderate activity, disease scores decreased or increased, respectively ($p<0.05$). Despite prior literature suggesting the benefits of physical activity[186], the long longitudinal survey timepoints used in this study preclude any causal inferences. It is not possible to know, for example, whether increased physical activity levels reduced disease outcomes, or whether patients experiencing reduced symptoms were better able to exercise. Future studies that solicit more time points for symptom data from patients will help us investigate this.

Implications for Health Care and Research

This research has several implications for the use of wearable devices and PROs for patients with IBDs. The use of real-world data to identify phenotypes with similar activity attributes could be leveraged to develop interventions that promote self-management and coping abilities, as these are important components in managing IBDs[217, 218]. Previous studies have used wearables to personalize behavioral coaching strategies, which resulted in improvements in physical activity and clinical biomarkers (e.g., lipids, hemoglobin A1c levels)[80, 81]. In our study, the low activity group was characterized by short durations of moderate to vigorous activity and low levels of exercise. Interventions effective in improving sedentary behavior often utilize established behavioral change techniques, including goal setting, self-monitoring, or social support[219]. We found the greatest difference among PROMIS domains between the low and high activity groups for social relationships, and evidence suggests social support may improve psychological symptoms and self-management behaviors in patients with IBDs[217, 220]. New models of patient-centered care have been proposed, such as the IBD specialty medical home, which involve multifaceted approaches focusing on social support, behavioral skills, and stress management techniques[221]. Providing outreach to patients with changes in health status or physical activity, which could be indicated by cluster transitions, could be an important aspect of personalized IBD care. Moreover, as new therapies are being developed, the ability to observe distinguishable physical activity phenotypes could reveal insights related to the impact of pharmacotherapies within clinical trials or at the point-of-care. However, if real-time data are used within clinical care or for just-in-time interventions, necessary protocols and validation strategies are needed to ensure accurate data are presented to patients and care teams in a meaningful and easily interpretable way[222]. Visual analytics and additional machine learning approaches are being developed and evaluated as part of the Precision VISSTA study.

Patient-generated health data also have potential to facilitate detection of inflammation. Evidence suggests certain inflammatory responses may be able to be detected through physiological measurements or lifestyle characteristics from wearables, such as elevated heart rate, elevated skin temperature, or sleep deviations[223, 224]. Very few studies have assessed the relationship between wearable physiological measurements and inflammatory responses in patients with IBDs[204, 205]. Sossenheimer et al (2019) found lower daily steps within the week prior to elevated inflammatory biomarkers (i.e., C-reactive protein and faecal calprotectin), but did not find differences in resting heart rate[204]. Wiestler et al (2019) also found lower levels of physical activity in patients with active disease compared to those in remission, in addition to lower sleep efficiency[205]. Given the vast amount of data from sensor devices, robust analytical pipelines are necessary to process, analyze, combine with other data streams, and derive actionable information from the data[225]. Unsupervised learning models, which do not require costly labeled data, are useful for partitioning large datasets into smaller groups of related information. These related subgroups have potential to inform strategies for detection or mitigation of inflammatory responses given the fluctuating inflammatory symptoms and disease trajectories that vary by patient.

<u>Limitations</u>

Participants in the CCFA Partners Internet-based cohort may not be representative of population-based IBD cohorts (e.g., higher proportion of women, higher educational attainment). Those who connected a smartphone app or wearable device to the portal represents a subset who have access to these devices, sufficient connectivity, and may already be motivated to track their health. Our sample was more physically active than the general population, taking approximately 7,900 steps per day. It is estimated the U.S. population takes 4,800 steps per day (~5,000 steps worldwide)[226]. Thus, our findings may not be representative of all patients with IBDs.

There are some limitations for using consumer-grade wearable device data. Most brands do not make the details of their algorithms or firmware updates available, so there may be differences in hardware or sensors across brands and devices over time. For example, moderate to vigorous active durations are defined by the type of device. Innate user differences may also exist, such as the location the device was worn (dominant vs. non-dominant hand), wear time, or accuracy of manually logged exercise, which cannot be verified in the existing data. Lastly, our clustering approach was limited by the lack of gradient or differentiation between individuals once clusters were established. Therefore, it is possible that individuals with similar activity patterns could be in different clusters. Clusters also represent a snapshot of physical activity patterns so we were not able to establish causal relationships.

Conclusion

Extracting patterns and changes in lifestyle behaviors in real-world usage of wearable devices to track physical activity and their association with PROs could inform personalized treatments and interventions for patients with IBDs. The deluge of data generated from patients necessitates innovative analytic pipelines to confer meaningful insights that promote long-term IBD remission. Unsupervised learning techniques may be valuable to cluster multidimensional lifestyle-related characteristics. Patients in the low physical activity cluster reported the worst psychosocial health and disease activity (i.e., depression, pain, fatigue, sleep disturbance, social relationships, and Crohn's Disease activity) compared to those in moderate and high activity clusters. Additional support for physical and psychosocial symptoms and exercise may be valuable for those low physical activity IBD subgroups. Future research should assess these findings across among more diverse cohorts and with more frequent prospective PROs measurement. Additional investigation should examine associations with additional lifestyle and treatment characteristics to detect flare-ups in disease activity, prevent exacerbation, or to develop potential interventions for patients with IBDs.

CHAPTER 5: SUMMARY AND CONCLUSION

Summary of Findings

Leveraging a user-centered design process, we designed, developed, and evaluated the usability of a hypertension self-management conversational agent ("Medicagent"). During the initial phase, patients' needs and perceptions towards using conversational agents to assist with managing their blood pressure and medication regimens were elicited (Chapter 2). Based on these patient-defined needs, a functional prototype of Medicagent was iteratively designed and developed using Google Cloud's Dialogflow natural language understanding engine. Patients then interacted with Medicagent during a task-based usability testing session where we evaluated how usable and acceptable it was to complete self-management tasks (Chapter 3). As physical activity is an important component of hypertension self-management, we also leveraged wearable device and app data that could be used to inform future contextually aware dialogues. At the time the work was conducted, no hypertension-related wearable device datasets were available due to COVID-19 restrictions, so we leveraged an existing dataset from patients with IBDs. Using an unsupervised learning approach, we identified physical activity clusters, examined how individual patients moved among the clusters longitudinally, and assessed the association between physical activity clusters and health status (Chapter 4). This approach could be used to inform tailored conversational agent coaching strategies based on changes in physical activity and applied to hypertension cohorts when datasets become available. A synopsis of our findings is described below.

In Chapter 2, qualitative analysis revealed that participants had generally positive attitudes and curiosity towards using a conversational agent for hypertension self-management. While most (80%) had not previously used a conversational agent, almost all perceived that the conversational nature would be helpful for various self-management tasks such as tracking medications, refills, communicating with care team members, and maintaining a healthy lifestyle. Most desired integration with patient portals, pharmacy apps, health tracking apps, or wearable devices. Several described how a conversational agent could provide feedback on results after a clinic visit or encouragement on health-related metrics (e.g., blood pressure, weight, physical activity). Many perceived conversational agents as humanlike and felt they could provide accountability to keep their blood pressure under control. Those who were younger and taking fewer medications were the most curious about using a conversational agent. Characteristics varied with participants' perspectives on how often they desired to interact with a conversational agent. Despite general enthusiasm, participants expressed several concerns with conversational agents providing too much information, sending demanding lifestyle change messages, invading their privacy, or having usability issues with interactions on a cell phone.

Based on user needs identified, a functional prototype of a text-based conversational agent ("Medicagent") was iteratively designed, developed, and evaluated (Chapter 3). During the usability evaluation, almost all (98%) tasks were successfully completed across ten participants in an average of 18 minutes. Participants finished data retrieval tasks faster on average (58 sec) as compared to data entry tasks (93 sec). The time spent completing tasks decreased with each additional data entry task, but varied with the data retrieval tasks. This may be due to the two data retrieval tasks that included hyperlinks to a website, in which participants reported difficulty navigating between browser tabs on the computer. Several described how there was a short learning curve for using Medicagent, and the general decrease in task duration suggests participants may have become more comfortable interacting as they became more familiar with the interface. We observed two main types of interactions styles using buttons and utterances, and several participants used a combination of both. In regards to the system performance, 8.6% (11/128) of utterances were unsuccessfully mapped to an intent. These errors were due to unrecognized spelling or formatting of dates, times, and blood pressure inputs. On average, participants rated Medicagent 78.8 on the System Usability Scale (SUS), which demonstrates acceptable usability[168, 169]. To our knowledge, there are no reported SUS scores of hypertension-related conversational agents, but mental health conversational agents have been reported as 81.8[71] and 88.2[170]. However, participant characteristics differed among studies, and our study had an older sample which might impact the scores. Some had difficulties navigating without standard app features like menu and back buttons, reported the need for a health professional persona, and felt the unrecognized utterance error messages were too generic. These are the main areas identified for refinement of Medicagent to prepare for future pilot testing and may facilitate usability and self-management tasks.

In Chapter 4, a total of 430 patients with IBDs were clustered into groups of low physical activity (33.7%), moderate physical activity (46.0%), and high physical activity (20.3%). Clusters were based on average number of steps, distance (miles), and moderate to vigorous active duration (minutes) over 6-week periods from patients' wearables or fitness apps. Overall, the clusters were moderately defined (average silhouette coefficient=0.54). The quality of the clustering was highest in the low activity cluster (silhouette coefficient=0.60), suggesting that participants had the most similarities in physical activity within this cluster. Sociodemographic characteristics varied across clusters, and those with low activity were older, had higher BMIs, and longer disease duration compared to the other clusters ($p<0.05$). Those in the low activity cluster also reported worse health status for depression, pain, fatigue, sleep disturbance, social relationships, and the short Crohn's Disease Activity Index compared to those with moderate and high physical activity ($p<0.01$). When we longitudinally assessed changes in physical activity and health status across 6-month periods that coincided with survey timepoints, exercise patterns largely did not fluctuate. Approximately 68% of patients remained in their original cluster, and only about 1% transitioned to or from the furthest clusters of low and high activity. This indicates exercise levels are largely stable. Conversational agent dialogues could serve as just-in-time adaptive interventions to provide meaningful information to patients before health status potentially declines or to increase physical activity. For example, a conversational agent could initiate communication with a patient based on changes in health status or physical activity, which could be indicated by cluster transitions. These results should be validated with other cohorts and with more frequent health survey measurements prior to developing such an intervention.

Implications for Health Care and Research

There are a number of implications for health care practice and research. First, our findings from the initial design stage have important considerations for the informational and support needs of conversational agents in patients with hypertension. Patients' self-management needs were consistent with prior research for other digital health approaches[27, 28, 38, 42, 101], and our research extends these needs by examining conversational components. The types of conversations patients desired varied, but the most common topics included encouragement on health-related metrics and feedback on results after a clinic visit. Some preferred to have a more active coaching style while others felt it might be too intrusive. These differences may be due to a variety of sociodemographic factors, clinical characteristics, behavioral determinants, or comfort with technology. Thus, it may be important for users to customize their preferred communication or coaching style of the conversational agent. For example, a more active coaching agent could provide frequent or didactive feedback as compared to a passive coach that employs sporadic check-ins. As demonstrated through the Computers are Social Actors paradigm, humans exhibit social relationships with computers[227]. Individuals have a natural tendency to interact with computers as if they were people, and computers are perceived as more likeable when they provide compliments or make jokes[58, 59]. This suggests that conversational agents should not only provide tailored self-management conversations, but also converse using relational and social dialogues. These relational conversations vary across cultures and communities, and thoughtful approaches to developing the knowledge base should be made to reduce algorithmic biases that favor one group of users[228]. Ongoing research and evaluation of conversational agents across diverse patient characteristics is vital to prevent these biases and mitigate potential technology and socioeconomic driven disparities.

Similarly, our findings have several implications for the personification and ethical use of systems with human characteristics. Several participants described conversational agents as ‘friendly’ or ‘like talking with a friend’ and desired for Medicagent to have visual and verbal characteristics of a health professional. Desired characteristics included a white coat, medical hat, and voice, and some participants felt these attributes would improve the credibility of the conversational agent. Embodied conversational agents that emulate face-to-face conversations using speech and nonverbal characteristics may be beneficial in establishing patient rapport and improving satisfaction[176]. Allowing users to customize anthropomorphic characteristics, such as race, gender, or dialect, may also remove design biases and improve user experience[229]. Nonetheless, there are number of ethical implications for communicating with an automated system that simulates a human. In Weizenbaum’s early experiments with ELIZA, users formed strong emotional attachment to ELIZA and even wanted to converse in private[230]. While users were aware they were communicating with a computer, they revealed sensitive information and often confided in the system[230]. Conversational agents may not be suitable for some patient populations, such as those with cognitive impairments or psychosis[231]. Although some patients may prefer to disclose sensitive information to a computer rather than a person, removing the human and lived experiences of individuals should not be substituted for computers[230, 231]. Providing options to communicate with a human prior to or while conversing with a conversational agent, appropriate disclosures, and transparency are necessary ethical components. Moral guidelines for users of such systems should also be considered to ensure ethically-aligned interactions.

Improving user interactions and patient safety for free-text input interactions of conversational agents is critical. We found that patients have diverse interaction preferences of conversational agents, and flexible interaction features may be encompassed through various visual cues (i.e., menu icons, buttons, text-input boxes) and multimodal interfaces. Users inputted a variety of formats for dates and blood pressure values, some of which were not recognized by Medicagent. Handling unrecognized errors gracefully is particularly important for conversational agents because people generally perceive robots as intelligent and competent[173, 174]. At this early stage with limited health care knowledge bases, it is unlikely a conversational agent would be able to recognize the vast number of possible free-text inputs. Therefore, adding context to error messages to enable users to better understand the cause of the error may reduce user frustration[175]. Incorporating features that may facilitate data entry, such as a calendar of dates, may also be valuable to minimize unrecognized utterances arising from data entry errors[155]. Despite the benefits of free-text inputs, these also pose risks to patient safety if the system misunderstands the input or lacks the ability to escalate urgent information provided by users. The World Economic Forum's Chatbots Reveal, Escalate, Substitute, Explain, and Track (RESET) project created a governance framework for the responsible use of chatbots in health care in 2020[232]. The framework outlines AI ethics principles and actions to operationalize them for different types of chatbots. They defined four types of chatbots according to the level of risk (i.e., low, moderate, high, very high) based on the International Medical Device Regulators Forum's software as a medical device document[233]. For example, low risk chatbots that assist with finding a doctor or scheduling an appointment require different oversight as compared to high risk chatbots that aid in diagnoses. Providing patient-friendly and regularly updated policies regarding the scope of use and limitations of the conversational agent may also help mitigate potential safety concerns.

Lastly, this research has several considerations for the use of wearable devices and PROs for patients with IBDs and other chronic diseases. The ability to observe patterns and changes in lifestyle behaviors throughout patients' daily lives supports opportunities to inform treatments and interventions. Many interventions to increase physical activity focus on step counts[234], though we considered additional metrics and patterns of use (i.e., distance, moderate to vigorous activity duration, calories burned, weekday use and weekend use). Clustering approaches are useful to combine multiple variables to more comprehensively understand these multifaceted physical activity patterns. Physical activity phenotypes derived from the clustering could be used to personalize strategies for detection or mitigation of inflammatory responses given the fluctuating inflammatory symptoms and disease trajectories that vary by patient. Wearable device data have been previously used to personalize behavioral coaching strategies, which resulted in improvements in physical activity and clinical biomarkers[80, 81]. Our findings begin to illustrate the physical activity patterns in patients with IBDs and their associations with health outcomes, although results should be validated before meaningful interventions can be developed. Recent work has focused on developing a conversational agent for patients with IBDs[180]. Zand et al (2020) used natural language processing to categorize electronic messaging data between patients with IBD and providers to generate a knowledge base for the chatbot[180]. Patients' questions and concerns were mapped to eight categories, such as symptoms, medications, appointments, and laboratory investigations[180]. Such approaches may be useful to develop the knowledge base for conversational agents for various chronic disease or multimorbid cohorts.

Directions for Future Research

There are a number of opportunities for future research involving the design, development, user experience, implementation, and efficacy of hypertension self-management and other health-related conversational agents. For Medicagent, our evaluation study revealed several refinements to improve user navigation, enhance error messages, enrich the health professional persona, and reduce data entry errors. As some participants desired voice interactions, additional research could compare the usability and user experience for self-management interventions via text, voice, and multimodal interfaces. Further, few public knowledge bases for health care conversational agents are available, and future work could develop and validate an open-source knowledge base related to self-management. Although this would involve a time-consuming health professional annotation process, the knowledge base could be valuable as self-management skills are similar across many chronic diseases, including medication adherence, physical activity, and healthy diet. The knowledge base could be used by designers and developers as an initial starting point and then tailored based on patients' needs and cultural factors.

As many health-focused conversational agents are in the early design phases and are not integrated within care delivery settings, development of a sociotechnical framework for integrating conversational agents within clinical care would be beneficial. The framework could build upon existing approaches that incorporate technical, workflow, and patient safety aspects of implementing AI tools and PGHD within health care. For technical aspects, the Health Level Seven International Fast Health care Interoperability Resources (FHIR) standard has become widely used to capture, integrate, and exchange information in EHRs[235]. A proof-of-concept conversational agent interoperability architecture was created to translate FHIR resources into Artificial Intelligence Markup Language[236]. This is useful because it has been widely to develop conversational agents. The proposed system allows the patient to upload an image or ask a question, store it in a FHIR database, and display the image or question to a provider[236]. Additional architectures for integrating PGHD into care settings that are already in use, such as SMART Markers[237], could also be built upon to incorporate linguistic data generated from conversational agents and then map the data to existing medical terminologies. For workflow and implementation processes, Li et al (2020) outlined the human and technical processes involved to safely and effectively implement AI in health care[238]. This approach involves multidisciplinary design, implementation, and evaluation using hybrid assessments of the system, efficacy, and integration[238]. Additional approaches such as the Systems Engineering Initiative for Patient Safety may also be beneficial in understanding the multilayered interactions between people, technologies, and workflow processes[239]. This model can be applied with both patients and providers in the center of the health care work system to identify areas that prevent patients from receiving high quality care and ensure technologies meet all stakeholder needs. Creating a sociotechnical framework for integrating conversational agents into health care settings also supports consideration of the ethical, liability, and privacy implications outlined above.

Similar to incorporating medical data within conversational agents, another direction for future research involves integrating wearables data to tailor the conversations. Prior studies have integrated apps and wearables with conversational agents[72, 76, 79], though the field lacks a comprehensive and reproducible framework. This research would involve development and validation of a model to deliver contextually aware conversational interventions from wearable device data. Variations across devices and malfunctions would also need to be incorporated to ensure safe and effective messaging. For example, a Food and Drug Administration approved blood pressure monitoring wearable device[240] and activity tracker could be paired with a conversational agent to help patients better understand how lifestyle behaviors impact their blood pressure in real-time. The conversational agent could move beyond summative statistics and generic motivation, which are typically provided by commercial wearable devices, to more personalized information based on changes in blood pressure or physical activity. Digital phenotyping algorithms, similar to our clustering approach, could identify similar subgroups and then trigger evidence-based blood pressure recommendations or tailored messages with health status changes. Although most conversational agents are still limited by the lack of ability to hold truly meaningful conversations and personalized interactions[87], developing such a framework may help to assess feasibility and utility of contextualizing conversational interventions.

Future investigations are also needed to evaluate the potential efficacy of conversational agents through pilot or more robust randomized controlled trials. Hypertension self-management conversational agents should be evaluated based on changes in patients' blood pressure values, medication adherence, and self-management practices. Few studies to date have evaluated the effectiveness of using conversational agents for hypertension self-management[109, 110]. These studies found improvements between the intervention and control groups on some aspects of self-management such as self-confidence, energy expenditure, and diet quality, but neither study found any differences in mean blood pressure or medication adherence[109, 110]. An additional outcome to examine with longer term use of conversational agents is the impact on the patient-provider relationship. The positive health outcomes and adherence associated with effective patient-provider interactions has been well-studied[241], but it is unclear how a conversational agent would impact this relationship as they also have potential to foster a therapeutic bond. Perhaps potential inconsistencies between the information provided by a conversational agent and health professional could lead to patient distrust in their care team[231]. Examining longer term patterns of use and efficacy may uncover potentially unintended consequences of health care conversational agents.

Conversational Agents During COVID-19 and Beyond

The COVID-19 pandemic has rapidly accelerated virtual care delivery and the adoption of digital health tools like conversational agents[242]. Health systems have faced high volumes of phone calls, patient portal messages, and appointment requests related to COVID-19[243]. A number of symptom assessment conversational agents have emerged to offset the demand and deliver curated information to patients or recommend triage based on their responses[243-249]. For example, the Centers for Disease Control and Prevention (CDC) deployed a COVID-19 symptom checker on their website[249], and the World Health Organization (WHO) launched a chatbot on Facebook Messenger and WhatsApp[248]. Many of these conversational agents have been deployed on publicly available health system websites[244-247] or within the patient portal to assist with self-triage and self-scheduling appointments[243]. Conversational agents could also provide remote symptom monitoring or longer-term self-management strategies for patients experiencing ongoing symptoms following a diagnosis of COVID-19[250, 251]. Although the majority use a text-based interface, some COVID-19 assessments have been deployed through multimodal or voice interfaces such as Apple's Siri[252] and Amazon's Alexa[253]. Overall, this suggests the utility of conversational agents to provide high quality information, triage, and support during public health crises. Additional investigation is needed to understand patients experience and the hand off to a human during the triage process.

As conversational agents are rapidly expanding across other industries, such as finance and e-commerce, individuals are likely becoming more familiar and comfortable with conversational interfaces. This increasing experience may also contribute towards greater adoption in health care, and we may see an increase in conversational agents for public health and health care operations beyond COVID-19. Early evidence suggests the availability of 24-hour access and operational benefits may be important for health system adoption[243]. For example, Babylon Health's conversational agent has already been integrated with hospital systems with the National Health Service in the United Kingdom[254]. Beyond self-management, conversational agents may offer a number of opportunities across the health care landscape for patients. Conversational agents could be valuable throughout their health care journey ranging from enrollment in benefits, health insurance support, finding a provider, providing directions or arranging transportation, appointment pre-screening, tracking symptoms, or post-discharge education[232]. Many of these use cases remain unexplored, so the feasibility and impact on care delivery are uncertain.

Conclusion

We focused on patients and their needs in designing, developing, and evaluating a conversational agent to support managing blood pressure and adhering to medication regimens. This research is important because ongoing self-management of hypertension-related behaviors, such as blood pressure monitoring, medications, and physical activity, is challenging yet essential to improve health outcomes. Overall, this research contributes towards informational and support needs, patient-centric design recommendations, and improving the usability of conversational agents for hypertension self-management. Our study also highlights unique considerations of using conversational interfaces for managing health, including coaching styles, tailored dialogues, interaction patterns, and the safety, ethical, regulatory, and privacy implications. We showed that clustering phenotypes of physical activity can be associated with positive associations between physical activity and health outcomes in patients with IBDs. The use of real-world wearable device data to identify physical activity phenotypes could be beneficial for tailoring these conversational interventions based on phenotypic changes. Within this emerging field, future opportunities for research are vast, and we describe several areas of investigation that may be beneficial to advance the understanding and value of conversational agents and connected devices. As the health care landscape continues to evolve, multidisciplinary research efforts are needed to thoughtfully study the ongoing impacts and potential consequences of conversational agents that support patients in managing their health.

APPENDIX A: SEMI-STRUCTURED INTERVIEW PROTOCOL

Recruitment

Do you take medication for high blood pressure?

We are conducting a research study to better understand how individuals manage their blood pressure medications and their perspectives towards using a chatbot. Participants will be invited to complete one visit to UNC to participate in either an interview session or testing session that will last approximately one hour.

You may be eligible if:

- You have been told by a doctor or health professional that you have high blood pressure (hypertension)
- You currently take one or more blood pressure medications
- Own a smartphone or tablet

To see if you are eligible for the study, please complete a brief questionnaire: <electronic questionnaire link>. If you are eligible, we will reach out to you soon. You will receive a $25 gift card as a small token of appreciation for your time after participating in the interview or testing session. Thank you for your consideration.

If you would like to be e-mailed the eligibility questionnaire above or have any questions, please contact Ashley at <e-mail>.

This project was determined to be exempt from federal human subjects research regulations.

Eligibility Questionnaire

1. What is your name?	______________________________
2. What is your age?	_____
3. What is your gender?	□ Female □ Male □ Other
4. What race do you consider yourself to be? One or more categories may be selected.	□ White or Caucasian □ Black or African American □ Asian or Asian American □ American Indian or Alaska Native □ Other, please specify: __________________________________
5. Do you consider yourself Hispanic or Latino or Latina?	□ Yes □ No
6. What is your highest level of education completed?	□ High school, GED, or less □ Some college □ College graduate or more
7. Has a doctor or health professional ever told you that you have high blood pressure (hypertension)?	□ Yes □ No
8. How many prescription medications are you currently taking on a daily basis? A prescription medication is one that requires a note from your doctor.	_____
9. How often do you use the Internet?	□ Less than once a week □ A few times each week □ Once a day □ More than once a day
10. Do you currently take one or more high blood pressure medications? 11. Are you able to take medications by yourself? 12. Do you own a smartphone? 13. Are you willing and able to provide informed consent?	□ Yes □ No

14. Is English your primary language for speaking and reading? 15. Do you have the ability to attend an in-person interview session in Chapel Hill, NC?	
Thank you for completing the survey. If you are eligible to participate, a member from the research team will contact you to schedule a one hour visit to UNC. Please provide your e-mail and phone number.	E-mail: ________________________ Phone: ________________________

Enrollment Questionnaire

Sociodemographics	
1. What is your age?	_____
2. What is your gender?	□ Female □ Male □ Other
3. What race do you consider yourself to be? One or more categories may be selected.	□ White or Caucasian □ Black or African American □ Asian or Asian American □ American Indian or Alaska Native □ Other, please specify: ____________________________________
4. Do you consider yourself Hispanic or Latino or Latina?	□ Yes □ No
5. What is your highest level of education completed?	□ Less than high school □ High school or GED □ Some college □ College graduate or more
6. What is your total household income in the past 12 months, before taxes, including money from a job, person, Social Security, or any other source?	□ Less than $20,000 □ $20,000 – $34,999 □ $35,000 – $49,999 □ $50,000 – $74,999 □ $75,000 – $99,999 □ $100,000 or more □ Do not wish to report
7. Including yourself, how many people are supported by this income?	_____
8. Has a doctor or health professional ever told you that you have one or more of the following conditions? One or more categories may be selected.	□ Arthritis □ Cancer □ Chronic obstructive pulmonary disease □ Diabetes □ Dementia □ Heart disease □ High blood pressure □ High cholesterol □ Other, please list all conditions: _______________________________
9. Approximately how long ago did a doctor or health professional	□ Less than 1 year ago □ 1 – 2 years ago

tell you that you have high blood pressure?	□ 3 – 5 years ago □ 5 or more years ago
10. How many prescription medications are you currently taking on a daily basis? A prescription medication is one that requires a note from your doctor.	_____
11. Which medications do you take for your blood pressure? Please list all.	____________________________________
12. Overall, how confident are you that your blood pressure is under control?	□ Completely confident □ Very confident □ Somewhat confident □ A little confident □ Not confident at all
Health Literacy (3-item literacy measure, Chew et al 2004[134])	
Medication Self-Efficacy (Self-efficacy for Managing Medications and Treatments, PROMIS Item Bank v1.0 2016[135]) – questions 16-23	
Barriers to Medication Adherence (Adherence Starts with Knowledge 12, Matza et al 2009[136]) – questions 24-35	
Technology Use	
36. How often do you use the Internet?	□ Less than once a week □ A few times each week □ Once a day □ More than once a day
37. Which of the following devices do you own? Please select all devices.	□ Smartphone □ Basic cellphone □ Tablet (iPad, Microsoft Surface, Amazon Fire, etc.) □ Computer (desktop or laptop) □ Other, please list all other devices: ____________________________________
38. Have you ever used a chatbot or virtual assistant? These are technologies that can communicate with you like you are chatting with a person but it's an automated tool with no real human chatting.	□ Yes, please list: _________________ □ No □ Don't know

Consent Form

University of North Carolina at Chapel Hill
Consent to Participate in a Research Study
Adult Participants
Consent Form Version Date: 9/10/2019
IRB Study # 19-2024
Title of Study: A Conversational Agent to Support Hypertension Medication Self-Management
Principal Investigator: Ashley Griffin
Principal Investigator Department: Carolina Health Informatics Program
Study Contact Telephone Number: <phone>
Study Contact E-mail: <e-mail>
Funding Source: Carolina Health Informatics Program

CONCISE SUMMARY

A study team at the University of North Carolina is conducting a research study to better understand how patients manage their blood pressure and medications as well as perspectives towards a chatbot. A chatbot is a system that can communicate with people. It is often called a virtual assistant or coach even though it’s not an actual person. We are looking for opinions and preferences from patients to help us with the design.

Participants will be invited to complete a one-on-one interview session that will last approximately one hour.

Participating in this research study is voluntary and you may choose not to participate, or you may withdraw your consent to be in the study, for any reason, without penalty.

<u>What are some general things you should know about research studies?</u>

You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty. Deciding not to be in the study, now or later, will not affect your ability to receive medical care at UNC or your employment/student status.

Research studies are designed to obtain new knowledge. This new information may help people in the future. Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study. You will be given a copy of this consent form. You should ask the researchers named above, or staff members who may assist them, any questions you have about this study at any time.

What is the purpose of this study?

The purpose of this research study is to better understand how patients manage their blood pressure and medications and perspectives with a chatbot. A chatbot is a system that can communicate with people. It is often called a virtual assistant or coach even though it's not an actual person. The results of this study will inform the development of a chatbot to support patients with managing their hypertension (high blood pressure) medications.

How many people will take part in this study?

If you decide to be in this study, you will be one of approximately 15 people in this research study.

How long will your part in this study last?

Your participation in a one-on-one interview session will last approximately one hour.

What will happen if you take part in the study?

You will participate in an individual interview session lasting approximately one hour with a member of our study team, where you will be asked about your current behaviors for managing blood pressure and medications as well as your perspectives and use of a chatbot. The session will be audio-taped so we can capture comments in a transcript for analysis.

What are the possible benefits from being in this study?

Research is designed to benefit society by gaining new knowledge. You may not benefit personally from being in this research study.

What are the possible risks or discomforts involved from being in this study?

We do not anticipate any risks or discomfort to you from being in this study. All comments made during the interview session will be kept confidential. Therefore, we encourage you to be as honest and open as you can.

How will information about you be protected?

Every effort will be taken to protect your identity as a participant in this study. You will not be identified in any report or publication of this study or its results. Your name will not appear on any transcripts; instead, you will be given a code number. Any reference to names in the audio files will be removed for analysis. The list which matches names and code numbers will be kept in a password protected electronic file kept on an encrypted, password-protected computer. After the interview session has been transcribed, the audio file will be destroyed after all analysis is completed. The audio recording may be requested to be turned off at any time during the interview.

Will you receive anything for being in this study?

You will receive a $25 gift card for taking part in this study.

Will it cost you anything to be in this study?

There will be no costs for being in the study except for your time.

What if you are a UNC student?

You may choose not to be in the study or to stop being in the study before it is over at any time. This will not affect your class standing or grades at UNC-Chapel Hill. You will not be offered or receive any special consideration if you take part in this research.

What if you are a UNC employee?

Taking part in this research is not a part of your University duties, and refusing will not affect your job. You will not be offered or receive any special job-related consideration if you take part in this research.

What if you have questions about this study?

You have the right to ask, and have answered, any questions you may have about this research. If you have questions, or concerns, you should contact the researchers listed on the first page of this form.

What if you have questions about your rights as a research participant?

All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject you may contact, anonymously if you wish, the Institutional Review Board at <phone> or by e-mail to <e-mail>.

Do you agree to be audio-taped during this interview session?

Yes___

No ___

Participant's Agreement:

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to participate in this research study.

____ ____________

Signature of Research Participant Date

____

Printed Name of Research Participant

Signature of Research Team Member Obtaining Consent Date

Printed Name of Research Team Member Obtaining Consent

Interview Guide

Welcome Introduction (15 minutes)

Hello. My name is Ashley and I'll be interviewing you today to learn about how you manage your blood pressure and medications and to see what your thoughts are on using chatbots (I'll go over what these are a bit later).

Before we get started, there is some information that we need to cover:

- Voluntary Participation: Your participation in the interview is entirely voluntary. You may stop participating at any time. You do not have to answer any questions that you do not wish to answer. You may withdraw from the interview at any time with no consequences. The consent forms provide more detailed information regarding confidentiality and the voluntary nature of participation, which I will review with you now. Please read and sign the consent form.

<Sign consent form>

- Confidentiality: Everything that you say here will be kept strictly confidential. Nothing said will be associated with your name.

Before we begin the interview, please complete this questionnaire.

<Administer Qualtrics questionnaire>

- Audio-Taping: This session is being taped so that I can write an accurate report of what was said and will be secured safely. Do I have your permission to record our session?

<Begin audio-taping>

Our interview will be divided into two parts. For the first part of the interview I will ask you questions about how you manage your blood pressure and medications, and for the second part I will ask you about your attitudes towards using a new technology to support your blood pressure and medications. Feel free to be honest and open about how you feel about the technology. There are no right or wrong answers and we really want to get your opinions (good or bad). Any questions before we start?

Part 1: Behavioral, motivational, and informational needs (20 minutes)

1. Can you tell me about what you do to manage or control your blood pressure?
    - Probe: What types of information or resources, if any, would be helpful for you to manage your blood pressure?

- Probe: Do you prefer visual or written information? Examples would be videos or infographics.
- Probe: What type of support may be helpful to you to manage your blood pressure?
    - Probe: Who are the people you get support from to manage your blood pressure? (Would support from a friend, family member, or health professional be helpful for you to manage your blood pressure?)

2. How do you keep track of taking your medications?
    - Probe: Can you tell me more about that?
    - Probe: What would help you with keeping track of taking your medications?
    - Probe: What types of information or resources, if any, would be helpful for you to track your medications?
    - Probe: What type of support do you feel would be helpful?
        - Probe: Who are the people who help you keep track of taking your medications? (Would support from a friend, family member, or health professional be helpful for you to track your medications?)

3. What about refills? How do you keep track of refilling your medications?
    - Probe: What would help you with refilling your medications?
    - Probe: What information be helpful for you to refill your medications?
    - Probe: What type of support do you feel would be helpful?
        - Probe: Who are the people who help you refill your medications? (Would support from a friend, family member, or health professional be helpful?)

4. Have you ever used technology to help you track your medications or refills? This could be a computer, tablet, or mobile app.
    - Yes: Can you tell me about how you've used a device or app to help track your medications or your refills?
    - No/don't know: Have you ever used a diary to track your medications or refills? (Would you be interested in using technology to track your medications or your refills?)

5. What, if anything, would help motivate you to take and refill your medications?
    - Probe: What types of reminders would you want to help you take your medications as prescribed daily?

Thank you so much! We're done with the first part. Is there anything you would like to add that you think I may have missed that you want to tell me about?

Part 2: Conversational agent (20 minutes)
For the second part of the interview, I will show you an example of a chatbot which is a system that can communicate with people. It is often called a virtual assistant or coach even though it's not an actual person. I will show you a short video of an example of a chatbot that helps people manage their health and wellness care. As you watch the video, imagine using one of these to help you manage your blood pressure and medications. After the video, I'll ask you for some of your thoughts related to chatbots.

<Play video: https://www.youtube.com/watch?time_continue=60&v=BtqJaHv53g0>

6. What do you think about the chatbot in the video? Would you use something like it?

7. Now imagine that you would use something similar to the chatbot for managing your blood pressure or medications and that you could use it on your smartphone. How do you think a chatbot could help you manage your blood pressure?
    - Probe: What types of things would you want to see in the chatbot?
    - Probe: What would make you want to use the chatbot?
    - What would get in the way of using a chatbot to help you take and refill your medications?

8. How do you think a chatbot could help you take or refill your medications?
    - Probe: For example, a chatbot could send you a reminder to take your medication at the right time. Would something like that be helpful?
    - Probe: Would it be useful if the chatbot helped you schedule appointments at a pharmacy or clinic?
    - Probe: Would you be interested in the chatbot providing you with tips or information to help you control your blood pressure?

9. How often would you want to interact with the chatbot every day?
    - Probe: Would you want to interact each day?
    - Probe: For what length of time each session?

10. Is there something you'd like to talk about that I haven't asked you?

Thank you for arranging your schedule today to be here for this session. We really appreciate you giving us your time, and opinions.
<Pay incentive and sign receipt form>
</End audio-taping>

APPENDIX B: USABILITY TESTING PROTOCOL

Recruitment

Do you take medication for high blood pressure?

We are conducting a research study to better understand how individuals manage their blood pressure medications and their perspectives towards using a chatbot. Participants will be invited to complete a remote testing session using UNC's secure and HIPAA-compliant Zoom video conferencing that will last approximately 1.5 hours.

You may be eligible if:

- You have been told by a doctor or health professional that you have high blood pressure (hypertension)
- You currently take one or more blood pressure medications
- Have access to a computer and able to use it without accessibility tools (e.g., you do not require special software such as screen readers or alternative controls)

To see if you are eligible for the study, please complete a brief questionnaire: <electronic questionnaire link>. If you are eligible, we will reach out to you soon. You will receive a $50 electronic gift card delivered to your e-mail address as a small token of appreciation for your time after participating in the testing session. Thank you for your consideration.

If you would like to be e-mailed the eligibility questionnaire above or have any questions, please contact Ashley at <e-mail>.

This study was approved by the UNC Office of Human Research Ethics Non-Biomedical Institutional Review Board (#19-2024) on 10/28/2019. The study was determined to be exempt from federal human subjects research regulations.

Eligibility Questionnaire

1. What is your name?	________________________________
2. What is your age?	_____
3. What is your gender?	□ Female □ Male □ Other
4. What race do you consider yourself to be? One or more categories may be selected.	□ White or Caucasian □ Black or African American □ Asian or Asian American □ American Indian or Alaska Native □ Other, please specify: ____________________________________
5. Do you consider yourself Hispanic or Latino or Latina?	□ Yes □ No
6. What is your highest level of education completed?	□ High school, GED, or less □ Some college □ College graduate or more
7. Has a doctor or health professional ever told you that you have high blood pressure (hypertension)?	□ Yes □ No
8. How many prescription medications are you currently taking on a daily basis? A prescription medication is one that requires a prescription from your doctor.	_____
9. How often do you use the Internet?	□ Almost constantly □ Several times a day □ About once a day □ Several times a week □ Less than once a week
10. Do you currently take one or more high blood pressure medications? 11. Are you able to take medications by yourself (without assistance)? 12. Are you willing and able to provide consent for yourself?	□ Yes □ No

13. Is English your primary language for speaking and reading? 14. Do you own or have access to a computer you could use during the study session? 15. Are you willing to allow us to make audio and screen recordings of the session? 16. Do you have a video or web camera on your computer? 17. Do you have a microphone on your computer? (Please mark ‘yes’ if you have a video or web camera.) 18. Are you able to use your computer without accessibility tools? (For example, you do not require special software such as screen readers or alternate controls.)	
Thank you for completing the survey. If you are eligible to participate, a member from the research team will contact you to schedule a remote session that will last 1.5 hours and an optional computer set-up session that will be approximately 20-30 minutes. Please provide your e-mail and phone number.	E-mail: ________________________ Phone: ________________________

Enrollment Questionnaire

Sociodemographics	
1. What is your age?	_____
2. What is your gender?	□ Female □ Male □ Other
3. What race do you consider yourself to be? One or more categories may be selected.	□ White or Caucasian □ Black or African American □ Asian or Asian American □ American Indian or Alaska Native □ Other, please specify: ____________________________________
4. Do you consider yourself Hispanic or Latino or Latina?	□ Yes □ No
5. What is your highest level of education completed?	□ Less than high school □ High school or GED □ Some college □ College graduate or more
6. What is your total household income in the past 12 months, before taxes, including money from a job, person, Social Security, or any other source?	□ Less than $20,000 □ $20,000 – $34,999 □ $35,000 – $49,999 □ $50,000 – $74,999 □ $75,000 – $99,999 □ $100,000 or more □ Do not wish to report
7. Including yourself, how many people are supported by this income?	_____
8. Has a doctor or health professional ever told you that you have one or more of the following conditions? One or more categories may be selected.	□ Arthritis □ Cancer □ Chronic obstructive pulmonary disease □ Diabetes □ Dementia □ Heart disease □ High blood pressure □ High cholesterol □ Other, please list all conditions: ____________________________________

9. Approximately how long ago did a doctor or health professional tell you that you have high blood pressure?	□ Less than 1 year ago □ 1 – 2 years ago □ 3 – 5 years ago □ 5 or more years ago
10. How many prescription medications are you currently taking on a daily basis? A prescription medication is one that requires a prescription from your doctor.	_____
11. Which medications do you take for your blood pressure? Please list all of them.	____________________________________
12. Overall, how confident are you that your blood pressure is under good control?	□ Completely confident □ Very confident □ Somewhat confident □ A little confident □ Not confident at all
Health Literacy (3-item literacy measure, Chew et al 2004[134])	
Medication Self-Efficacy (Self-efficacy for Managing Medications and Treatments, PROMIS Item Bank v1.0 2016[135]) – questions 16-23	
Barriers to Medication Adherence (Adherence Starts with Knowledge 12, Matza et al 2009[136]) – questions 24-35	
Technology Use	
36. How often do you use the Internet?	□ Almost constantly □ Several times a day □ About once a day □ Several times a week □ Less than once a week
37. Which of the following devices do you own? Please select all devices.	□ Smartphone □ Basic cellphone □ Tablet (iPad, Microsoft Surface, Amazon Fire, etc.) □ Computer (desktop or laptop) □ Other, please list all other devices: ____________________________________

38. Have you ever used a chatbot or virtual assistant? These are systems you can chat with but it's an automated tool with no real human chatting. For example, you could chat with one on a website or app, or you could speak to one like Apple Siri or Amazon Alexa.	□ Yes, please list: ___________________ □ No □ Don't know

System Usability Scale (Brooke et al 1996[167])

Consent Form

University of North Carolina at Chapel Hill
Consent to Participate in a Research Study
Adult Participants
Consent Form Version Date: 7/29/2020
IRB Study # 19-2024
Title of Study: A Conversational Agent to Support Hypertension Medication Self-Management
Principal Investigator: Ashley Griffin
Principal Investigator Department: Carolina Health Informatics Program
Study Contact E-mail: <e-mail>
Funding Source: Carolina Health Informatics Program

CONCISE SUMMARY

A study team at the University of North Carolina is conducting a research study to better understand how patients manage their blood pressure and medications as well as perspectives towards a chatbot. A chatbot is a system that can communicate with people. It is often called a virtual assistant or coach even though it's not an actual person. We are looking for opinions and preferences from patients to help us with the design.

Participants will be invited to complete a one-on-one remote testing session of the chatbot using a secure, HIPAA-compliant version of Zoom video conferencing that will last approximately 1.5 hours.

Participating in this research study is voluntary and you may choose not to participate, or you may withdraw your consent to be in the study, for any reason, without penalty.

<u>What are some general things you should know about research studies?</u>

You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty. Deciding not to be in the study, now or later, will not affect your ability to receive medical care at UNC or your employment/student status.

Research studies are designed to obtain new knowledge. This new information may help people in the future. Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study. You should ask the researchers named above, or staff members who may assist them, any questions you have about this study at any time.

What is the purpose of this study?

The purpose of this research study is to better understand how patients interact with a chatbot and perspectives of it. A chatbot is a system that can communicate with people. It is often called a virtual assistant or coach even though it's not an actual person. The results of this study will inform modifications to the chatbot to support patients with managing their hypertension (high blood pressure) medications.

How many people will take part in this study?

If you decide to be in this study, you will be one of approximately 10 people in this research study.

How long will your part in this study last?

Your participation in a one-on-one testing virtual session will last approximately 1.5 hours. You can optionally participate in a help session before your testing session.

What will happen if you take part in the study?

You will participate in a remote testing session using UNC Zoom with a member of our study team, where you will complete tasks one at a time using the chatbot. You will also be asked about your experience using the chatbot so that we can improve it. Your computer screen, audio, and video will be recorded, and the tasks will be timed so we can capture comments and metrics in a transcript for analysis.

What are the possible benefits from being in this study?

Research is designed to benefit society by gaining new knowledge. You may not benefit personally from being in this research study.

What are the possible risks or discomforts involved from being in this study?

We do not anticipate any risks or discomfort to you from being in this study. All comments made during this session will be kept confidential. Therefore, we encourage you to be as honest and open as you can. If at any point you feel that you are unable to complete a task, you may move to the next task.

How will information about you be protected?

Every effort will be taken to protect your identity as a participant in this study. You will not be identified in any report or publication of this study or its results. Your name will not appear on any transcripts; instead, you will be given a code number. The list which matches names and code numbers and recording files will be kept in a password protected electronic file kept on an encrypted, password-protected computer. After the session has been transcribed, the recordings will be destroyed after all analysis is completed. The screen, audio, or video recording may be requested to be turned off at any time during the session.

Will you receive anything for being in this study?

You will receive a $50 gift card for taking part in this study.

Will it cost you anything to be in this study?

There will be no costs for being in the study except for your time.

What if you are a UNC student?

You may choose not to be in the study or to stop being in the study before it is over at any time. This will not affect your class standing or grades at UNC-Chapel Hill. You will not be offered or receive any special consideration if you take part in this research.

What if you are a UNC employee?

Taking part in this research is not a part of your University duties, and refusing will not affect your job. You will not be offered or receive any special job-related consideration if you take part in this research.

What if you have questions about this study?

You have the right to ask, and have answered, any questions you may have about this research. If you have questions, or concerns, you should contact the researchers listed on the first page of this form.

What if you have questions about your rights as a research participant?

All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject you may contact, anonymously if you wish, the Institutional Review Board at <phone> or by e-mail to <e-mail>

Usability Testing Guide

Welcome Introduction (15 minutes)

Hello. My name is Ashley and I will be guiding you today as you interact with a chatbot, which is a tool that can communicate with you like you are chatting with an online person but it's an automated tool with no real human chatting. You may have heard of them referred to as virtual assistants or coaches.

Before we get started, there is some information that we need to cover:

- Your participation in the session is entirely voluntary. You may stop at any time, and you do not have to answer any questions that you do not wish to answer. You may withdraw from the session at any time with no consequences. You will need to complete the full session to receive the gift card. Everything that you say here will be kept strictly confidential. Nothing said will be associated with your name.
- The consent form provides more detailed information regarding confidentiality and the voluntary nature of participation. A copy the consent form was sent to you via e-mail. We will review this together now on my computer screen. <review with participant while sharing my screen>
- What questions do you have for me?
- Obtain verbal consent
    - Do you voluntarily agree to participate in this research study?
    - (If yes) Do you agree to have your screen, video, and audio recorded during this session?

<Allow time to complete Qualtrics questionnaire if the participant did not complete ahead of session>

Our session will be divided into two parts today. For the first part, I will introduce you to the chatbot and you'll complete short tasks. For the second part, I will ask you about your experience using the chatbot today. Please feel free to be completely honest and open with your feedback. It's the only way we can improve things and there are no right or wrong answers.

Part 1: Usability Testing (35 minutes)

You will complete tasks one at a time using the chatbot. For each task, I'd like for you to tell me about the steps you are taking and what you are thinking about as you are trying to do the tasks. You can talk about what you see and anything you find interesting or easy or difficult to do. I will show you a short video of what it means to think aloud now because it may feel a bit strange at first to talk through each of the steps you are taking while doing these tasks.

<Show example of thinking aloud by sharing my computer screen>

https://www.nngroup.com/articles/thinking-aloud-demo-video/

Do you have any questions about thinking aloud?

Once you feel that you've completed the task, please let me know. If at any point you feel that you are unable to complete the task, please also let me know so that we can move to the next task.

Now I would like for you to share your screen.

- I will send you a message in the chat window now that contains a website to access to chatbot.

I'll now tell you a little bit about the chatbot you'll be using today, which is called Medicagent. Medicagent is a virtual medication assistant focused on hypertension. The virtual assistant can provide you with information and help you keep track of your medications, refills, or blood pressure. It could also help you schedule appointments with your doctor, send you appointment reminders, or provide some health coaching. It could be accessed on your phone, tablet, or computer. Today we'll be using it on your computer. The pictures and information that you see in Medicagent are for demonstrative purposes during our testing session today.

May I proceed with the screen and video recording now?

<Begin recording to the cloud>

We will start with an example task. You can use the keyboard to chat or use the buttons that you see. Remember to think aloud as you are doing the task. Go ahead and click on the example task.

Now that the example task is completed, click on task #1.

--

You've just completed the first part of our session today. Please complete this questionnaire with the link provided in the chat. Try to record your gut reaction to each item, and don't think too much about items for a long time. Please check all items. If you feel that you cannot answer any of the questions, mark the number 3 in the center. <Send link to Qualtrics System Usability Scale in the chat window>

<u>Part 2: User Acceptability Semi-Structured Interview Guide (25 minutes)</u>

For the second part of this session, I'll ask you about what you thought of Medicagent. Even though you used it on your computer today during our session, you could also use it on your phone as an app if you preferred.

1. Overall what was your impression of using the chatbot Medicagent?

2. If you were to describe this chatbot to someone who has never seen it, what would you tell them?

3. What would have made this a better experience for you?

4. How do you think the chatbot could help you manage your blood pressure?
    - Probe: How do you think the chatbot could help you track your medications or refills?

5. What would you like to change about the chatbot?
    - Probe: What things are missing that you would like to add or take away?

6. What are your thoughts about the chatbot being able to connect to a device you might have, such as a blood pressure cuff, step tracker, sleep tracker, patient portal, or an app on your phone? (may want to ask about these one by one)
    - Probe: How would you feel if the chatbot could use that data to provide you with personalized tips or coaching?

7. What are your thoughts about the chatbot being connected to your pharmacy to help keep track of medication refills, expiration dates, or be able to order refills?
    - Probe: What other pharmacy information would be useful?

8. How do you feel about the chatbot being connected to your doctor's office to help schedule appointments or share information with your care team?
    - Probe: How do you think the chatbot could help you communicate with your care team?
    - Probe: What type of information would you want to share with your care team? (blood pressure, physical activity, sleep)
    - Probe: How useful would the chatbot be to help you and your doctor monitor your response to new medications or treatments?

9. What concerns do you have about the chatbot being connected to your medications or other medical information?

10. How often would you want to interact with chatbot?
    - Probe: Would you want to interact each day? Would this different depending on the circumstance?
    - Would the chatbot be more or less useful to you at certain times?

11. Would you want to use the chatbot for something other than blood pressure or medications?

- How helpful would a chatbot or tool like this one be to monitor symptoms and share them with the care team if you or a loved one were sick?
- How helpful would a chatbot be to help you prepare for a clinic visit or scheduled surgery?
- How useful would it be to follow-up with you or a loved one after the clinic or hospital visit?

This has been very helpful. Anything I missed and you'd like to share or provide any feedback?

Thank you for arranging your schedule today to be here for this session. We really appreciate you giving us your time, and opinions. You will receive an electronic gift card at your preferred e-mail.

<confirm e-mail address> Please check your junk mail in case it is there. I will be able to track whether you have received it.

</End recording>

Zoom Set-up Session

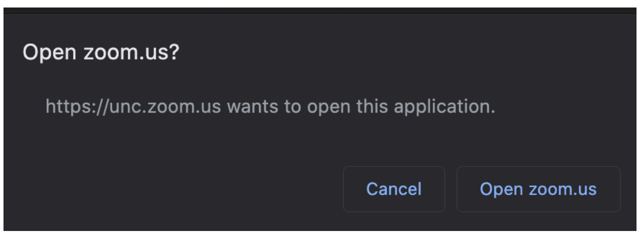
What is Zoom?

Zoom is an easy to use video and audio-conferencing service. It is free to use through a browser or the Zoom app on your computer and we will be using a special version that is secure and HIPAA-compliant to keep our session private.

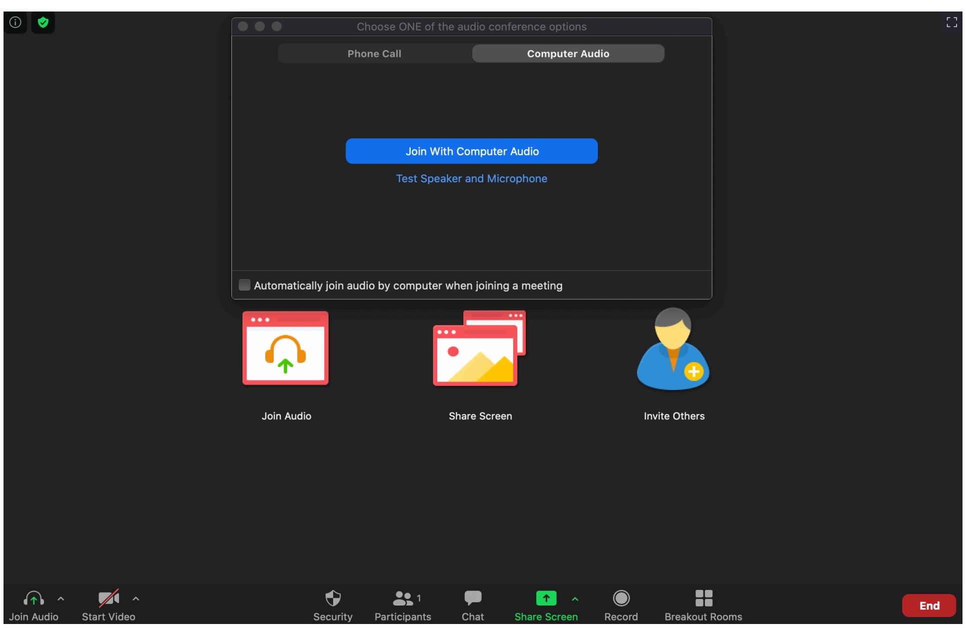
Joining a Zoom meeting

1. When it is time for you to join the session, please use your computer to click on this link: <zoom link>
    - If you are prompted for a meeting ID, enter: <meeting ID>

2. The link will open in a browser and you will see a dialog prompt. Click on Open zoom.us.



3. You will then enter the meeting room. Click on Join with Computer Audio.

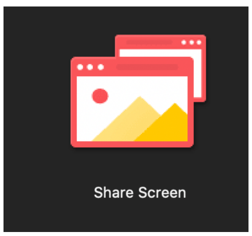


4. If you have a video or web camera on your computer, please click on Start Video in the bottom left.

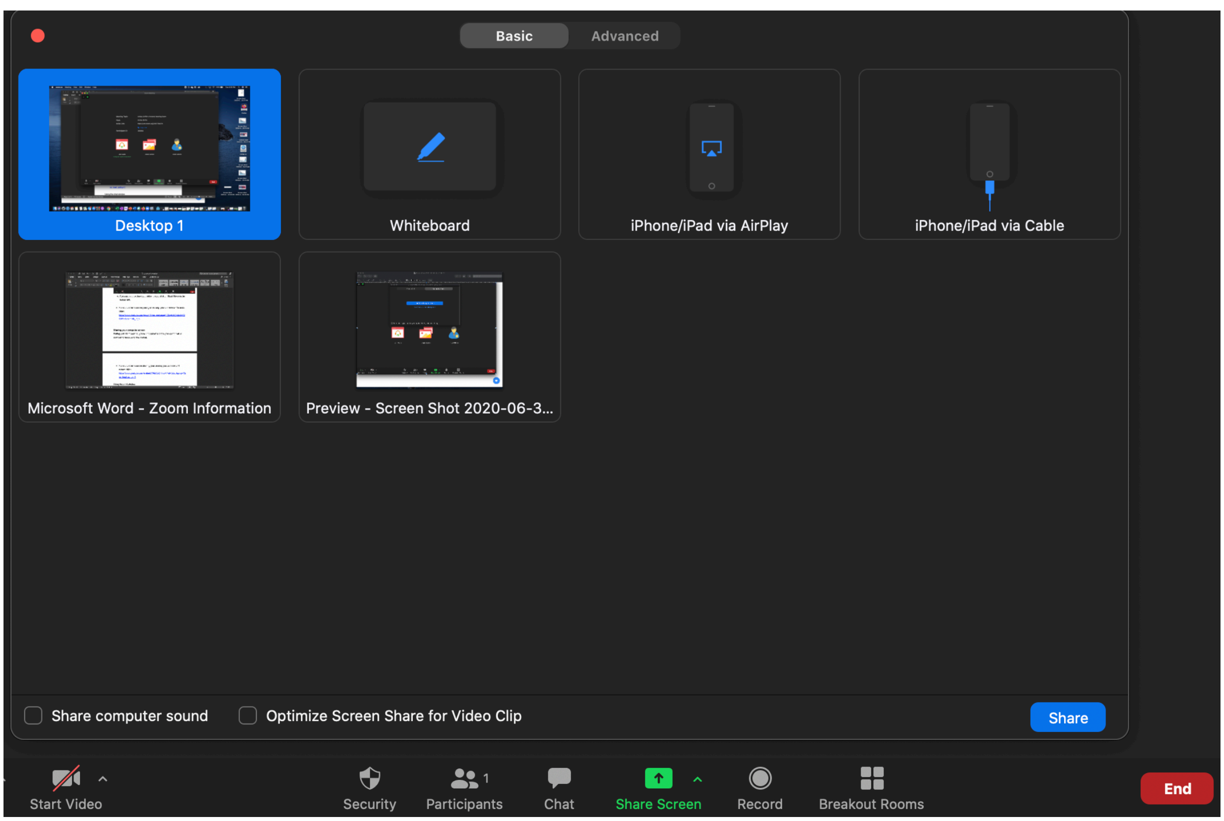
Sharing your computer screen

During part of the session, you will be asked to share your screen while completing tasks using the chatbot.

5. Please click on Share Screen.



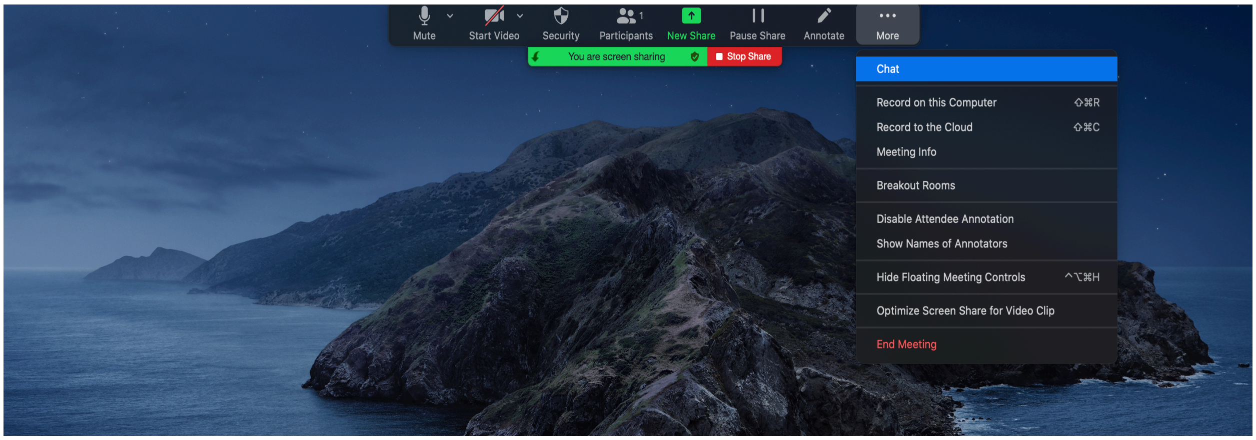
6. You will be prompted with different screens on your computer you can share. Click on Desktop, which allows you to share the contents of your desktop.



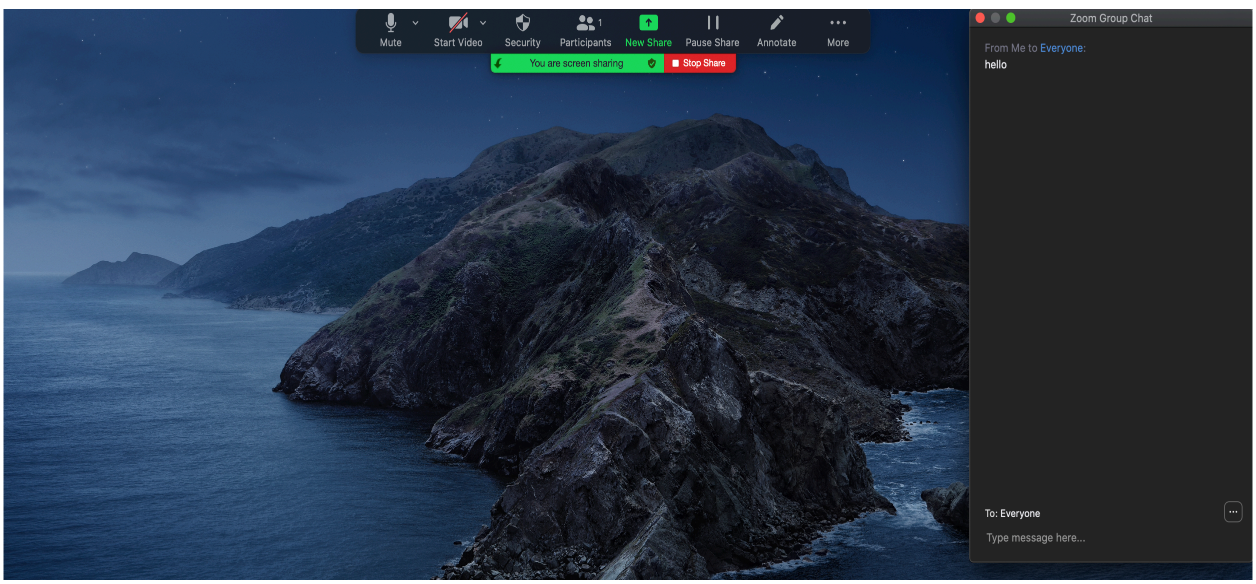
Using the chat window

During the session, you will receive links to questionnaires and tasks in the chat window.

7. Once you've started sharing your screen, click on More and then Chat to display the chat window.



8. Place the chat window on the right side of your screen. The left side of your screen will be used to interact with the chatbot which will provided to you during the session.



Additional information

- ~1 minute video about joining a Zoom meeting: https://www.youtube.com/watch?time_continue=10&v=hIkCmbvAHQQ&feature=emb_logo
- ~1 minute video about sharing your screen in Zoom: https://www.youtube.com/embed/YA6SGQlVmcA?rel=0&autoplay=1&cc_load_policy=1

REFERENCES

1. Centers for Medicare & Medicaid Services. National health expenditures highlights. 2017. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/highlights.pdf (22 May 2019).

2. Buttorff C, Ruder T, Bauman M. Multiple chronic conditions in the United States. 2017. https://www.rand.org/pubs/tools/TL221.html (1 May 2019).

3. Gerteis J, Izrael D, Deitz D, et al. Multiple chronic conditions chartbook. Agency for Healthcare Research and Quality. 2014. https://www.ahrq.gov/professionals/prevention-chronic-care/decision/mcc/resources.html (1 May 2019).

4. Centers for Disease Control and Prevention. About chronic diseases. 2020. https://www.cdc.gov/chronicdisease/about/index.htm#risks (15 Jan 2021).

5. Whelton PK, Carey RM, Aronow WS, et al. 2017 Guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension. 2018;71(6):1269-1324.

6. Barlow J, Wright C, Sheasby J, et al. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns. 2002;48(2):177-87.

7. Lorig K, Holman H, Sobel D, et al. Living a Healthy Life with Chronic Conditions: For Ongoing Physical and Mental Health Conditions. 4th ed. Boulder, CO: Bull Publishing Company 2013.

8. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med. 2003;26(1):1-7.

9. World Health Organization. Adherence to long-term therapies: evidence for action. 2003. http://apps.who.int/medicinedocs/en/d/Js4883e/ (15 May 2019).

10. Estacio EV, Whittle R, Protheroe J. The digital divide: Examining socio-demographic factors associated with health literacy, access and use of internet to seek health information. J Health Psychol. 2017;12(9):1668-1675.

11. Griffin AC, Chung AE. Health tracking and information sharing in the patient-centered era: a Health Information National Trends Survey (HINTS) study. AMIA Annu Symp Proc 2019:1041-1050.

12. Burnier M, Egan BM. Adherence in hypertension. Circulation Research. 2019;124(7):1124-40.

13. James PA, Oparil S, Carter BL, et al. 2014 Evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the eighth joint national committee. JAMA. 2014;311(5):507-520.

14. Orom H, Underwood W, Cheng Z, et al. Relationships as medicine: quality of the physician-patient relationship determines physician influence on treatment recommendation adherence. Health Serv Res. 2018;53(1):580-596.

15. Norris SL, Engelgau MM, Narayan KM. Effectiveness of self-management training in type 2 diabetes: a systematic review of randomized controlled trials. Diabetes Care. 2001;24(3):561-87.

16. Mullen PD, Green LW, Persinger GS. Clinical trials of patient education for chronic conditions: a comparative meta-analysis of intervention types. Prev Med. 1985;14(6): 753-81.

17. Schroeder K, Fahey T, Ebrahim S. Interventions for improving adherence to treatment in patients with high blood pressure in ambulatory settings. Cochrane Database of Systematic Reviews. 2004; (2):CD004804.

18. Chodosh J, Morton SC, Mojica W, et al. Meta-analysis: chronic disease self-management programs for older adults. Ann Intern Med. 2005;143(6):27-38.

19. Franek J. Self-management support interventions for persons with chronic disease: an evidence-based analysis. Ont Health Technol Assess Ser. 2013;13(9):1-60.

20. World Health Organization. mHealth: new horizons for health through mobile technologies. 2011. https://www.who.int/goe/publications/goe_mhealth_web.pdf. (20 Jun 2019).

21. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med. 2015;7(283):283rv3.

22. Pew Research Center. Mobile fact sheet. 2019. https://www.pewinternet.org/fact-sheet/mobile/ (2 July 2019).

23. Shapiro M, Johnston D, Wald J, et al. Patient-generated health data. White paper. RTI International. Prepared for Office of Policy and Planning, Office of the National Coordinator for Health Information Technology. 2012. https://www.rti.org/publication/patient-generated-health-data-white-paper (5 Jun 2019).

24. Chung AE, Sandler RS, Long MD, et al. Harnessing person-generated health data to accelerate patient-centered outcomes research: the Crohn's and Colitis Foundation of America PCORnet Patient Powered Research Network (CCFA Partners). J Am Med Inform Assoc. 2016;23(3):485-90.

25. Chung AE, Vu MB, Myers K, et al. Crohn's and Colitis Foundation of America Partners Patient-Powered Research Network: patient perspectives on facilitators and barriers to building an impactful patient-powered research network. Med Care. 2018; 56 (10 Suppl 1):S33-S40.

26. Alessa T, Abdi S, Hawley MS, et al. Mobile apps to support the self-management of hypertension: systematic review of effectiveness, usability, and user satisfaction. JMIR Mhealth Uhealth. 2018;6(7):e10723.

27. Rehman H, Kamal AK, Morris PB, et al. Mobile health (mHealth) technology for the management of hypertension and hyperlipidemia: slow start but loads of potential. Current Atherosclerosis Reports. 2017;19(3):12.

28. Ahmed I, Ahmad NS, Ali S, et al. Medication adherence apps: review and content analysis. JMIR Mhealth and Uhealth. 2018;6(3):e62.

29. Garabedian LF, Ross-Degnan D, Wharam JF. Mobile phone and smartphone technologies for diabetes care and self-management. Curr Diab Rep. 2015;15(12):109.

30. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res.2016;18(5):e97.

31. Sun R, Korytkowski MT, Sereika SM, et al. Patient portal use in diabetes management: literature review. JMIR Diabetes. 2018;3(4):e11199.

32. Jimison H, Gorman P, Woods W, et al. Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. Evid Rep Technol Assess (Full Rep). 2008(175):1-1422.

33. Ringeval M, Wagner G, Denford J, et al. Fitbit-based interventions for healthy lifestyle outcomes: systematic review and meta-analysis. J Med Internet Res. 2020; 22(10):e23954.

34. Cheatham SW, Stull KR, Fantigrassi M, et al. The efficacy of wearable activity tracking technology as part of a weight loss program: a systematic review. J Sports Med Phys Fitness. 2018;58(4):534-548.

35. Seto W, Leonard KJ, Cafazzo JA, et al. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res. 2012;14(1):e31.

36. Gokalp H, de Folter J, Verma V, et al. Integrated telehealth and telecare for monitoring frail elderly with chronic disease. Telemed J E Health. 2018;24(12):940-957.

37. Hamine S, Gerth-Guyette E, Faulx D, et al. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res. 2015;17(2):e52.

38. Thakkar J, Kurup R, Laba TL, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med. 2016;176(3):340-9.

39. Whittaker R, McRobbie H, Bullen C, et al. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016;4:CD006611.

40. Williams AD. Use of a text messaging program to promote adherence to daily physical activity guidelines: a review of the literature. Bariatric Nursing and Surgical Patient Care. 2012;7(1).

41. Arambepola C, Ricci-Cabello I, Manikavasagam P, et al. The impact of automated brief messages promoting lifestyle changes delivered via mobile devices to people with type 2 diabetes: a systematic literature review and meta-analysis of controlled trials. J Med Internet Res. 2016;18(4):e86.

42. Xiong S, Berkhouse H, Schooler M, et al. Effectiveness of mHealth interventions in improving medication adherence among people with hypertension: a systematic review. Current Hypertension Reports. 2018;20(10):86.

43. Logan AG, Irvine MJ, McIsaac WJ, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension. 2012;60(1): 51-57.

44. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc. 2018;25(9):1248-1258.

45. Jurafsky D, Martin J. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. 2nd ed. Upper Saddle River, NJ: Pearson Prentice Hall 2009.

46. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns. 2005;59(1):21-30.

47. Weizenbaum J. ELIZA – a computer program for the study of natural language communication between man and machine. Communications of the ACM. 1966;9(1): 36-45.

48. Rogers C. Client-centered therapy; its current practice, implications, and theory. Oxford, England: Houghton Mifflin 1951.

49. Amazon Web Services. Amazon Lex – Build Conversation Bots. https://aws.amazon.com/lex/ (15 Jun 2019).

50. Google Cloud. Dialogflow. https://dialogflow.com/ (19 Dec 2018).

51. Microsoft. Bot Framework. https://dev.botframework.com/ (19 Dec 2018).

52. Chung AE, Griffin AC, Selezneva D, et al. Health and fitness apps for hands-free voice-activated assistants: content analysis. JMIR Mhealth Uhealth. 2018;6(9):e174.

53. Bickmore T, Giorgino T. Health dialog systems for patients and consumers. Journal of Biomedical Informatics. 2006;39(5):556-571.

54. Klein J, Moon Y, Picard RW. This computer responds to user frustration: theory, design, and results. Interacting with Computers. 2002;14(2):119-140.

55. Bickmore TW, Silliman RA, Nelson K, et al. A randomized controlled trial of an automated exercise coach for older adults. J Am Geriatr Soc. 2013. 61(10):1676-83.

56. Bickmore TW, Schulman D, Sidner C. Automated interventions for multiple health behaviors using conversational agents. Patient Educ Couns. 2013;92(2):142-8.

57. Miner A, Chow A, Adler S, et al. Conversational agents and mental health: theory-informed assessment of language and affect. in Proceedings of the Fourth International Conference on Human Agent Interaction. 2016. Biopolis, Singapore.

58. Reeves B, Nass C. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. New York, NY: Cambridge University Press 1996.

59. Morkes J, Kernal HK, Nass C. Humor in task-oriented computer-mediated communication and human-computer interaction. CHI 98 Conference Summary on Human Factors in Computing Systems. 1998.

60. Guy I. Searching by talking: analysis of voice queries on mobile web search. Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2016.

61. Griffin AC, Zing X, Khairat S, et al. Conversational agents for chronic disease self-management: a systematic review. AMIA Annu Symp Proc. 2020. (In Press).

62. Centers for Medicare & Medicaid Services. Chronic Conditions. 2017. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Chronic-Conditions/CC_Main.html (19 Dec 2018).

63. International Standard Organization. Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: guidance on usability. 9241-11.1998. https://www.iso.org/standard/16883.html (20 Jan 2021).

64. Baskar J, Lindgren H. Human-agent dialogues on health topics - an evaluation study. Communications in Computer and Information Science. 2015;524:28-39.

65. Elmasri D, Maeder A. A conversational agent for an online mental health intervention. Brain Informatics and Health: International Conference 2016;243-251.

66. Fitzpatrick K, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health. 2017;4(2):e19.

67. Gaffney H, Mansell W, Edwards R, Wright J. Manage Your Life Online (MYLO): a pilot trial of a conversational computer-based intervention for problem solving in a student sample. Behav Cogn Psychother. 2014;42(6):731-46.

68. Kazemi DM, Cochran AR, Kelly JF, et al. Integrating mHealth mobile applications to reduce high risk drinking among underage students. Health Education Journal. 2013;73(3).

69. Kazemi DM, Borsari B, Levine MJ, Lamberson KA, Dooley B. REMIT: Development of a mHealth theory-based intervention to decrease heavy episodic drinking among college students. Addiction Research and Theory. 2018;26(5):377-385.

70. Ly KH, Ly A, Andersson G. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. Internet Interv. 2017; 10:39-46.

71. Schroeder J, Wilkes C, Rowan K, et al. Pocket skills: a conversational mobile web app to support dialectical behavioral therapy. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018;1-15.

72. Stein N, Brooks K. A fully automated conversational artificial intelligence for weight loss: longitudinal observational study among overweight and obese adults. JMIR Diabetes. 2017;2(2):e28.

73. Tsiourti C, Joly E, Wings C, et al. Virtual assistive companions for older adults: qualitative field study and design implications. Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare. 2014;57-64.

74. van Heerden A, Ntinga X, Vilakazi K. The potential of conversational agents to provide a rapid HIV counseling and testing services. International Conference on the Frontiers and Advances in Data Science. 2017.

75. Wang H, Zhang Q, Ip M, et al. Social media-based conversational agents for health management and interventions. Computer. 2018;51(8):26-33.

76. Watson A, Bickmore T, Cange A, et al. An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial. J Med Internet Res. 2012;14(1):e1.

77. Scholten MR, Kelders SM, van Gemert-Pijnen JE. Self-guided web-based interventions: scoping review on user needs and the potential of embodied conversational agents to address them. J Med Internet Res. 2017;19(11):e383.

78. Provoost S, Lau HM, Ruwaard J, et al. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res. 2017;19(5):e151.

79. Kramer LL, ter Stal S, Mulder BC, et al. Developing embodied conversational agents for coaching people in a healthy lifestyle: scoping review. J Med Internet Res. 2020;22(2):e14058.

80. Price ND, Magis AT, Earls JC, et al. A wellness study of 108 individuals using personal, dense, dynamic data clouds. Nat Biotechnol. 2017;35(8):747-756.

81. Hoye KV, Boen F, Lefevre J. The impact of different degrees of feedback on physical activity levels: a 4-week intervention study. Int J Environ Res Public Health. 2015;12(6):6561-6581.

82. Hardeman W, Houghton J, Lane K, et al. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act. 2019;16(1):31.

83. Schembre SM, Lio Y, Robertson MC, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. J Med Internet Res. 2018;20(3):e106.

84. Spruijt-Metz D, Wen CKF, O'Rielly, et al. Innovations in the use of interactive technology to support weight management. Curr Obes Rep. 2015;4(4):510-9.

85. Naughton F. Delivering "just-in-time" smoking cessation support via mobile phones: current knowledge and future directions. Nicotine & Tobacco Research. 2017;19(3)379-383.

86. Torous J, King MV, Lorme J, et al. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health. 2016;3(2):e16.

87. McTear M. Conversational AI: dialogue systems, conversational agents, and chatbots. San Rafael, CA: Morgan & Claypool Publishers 2020.

88. Bickmore TW, Trinh H, Olafsson S, et al. Patient and consumer safety risks when using conversational assistants for medical information: an observational study of Siri, Alexa, and Google Assistant. J Med Internet Res. 2018;20(9):e11510.

89. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. J Med Internet Res Ment Health. 2017;4(2):e19

90. Khatri AV, Hedayatnia B, Gabriel R, et al. Alexa Prize – State of the art in conversational AI. AI Magazine. 2018;39(3):40-55.

91. McGreevey JD, Hanson CW, Koppel R. Clinical, legal, and ethical aspects of artificial intelligence-assisted conversational agents in health care. JAMA. 2020;324(6): 552-553.

92. Google Cloud. Google Cloud services that are in scope for HIPAA. 2020. https://cloud.google.com/security/compliance/hipaa-compliances (20 Dec 2020).

93. Amazon Web Services. Voice for Healthcare. https://developer.amazon.com/en-US/alexa/alexa-skills-kit/get-deeper/custom-skills/healthcare-skills (20 Dec 2020).

94. Amazon Web Services. Introducing New Alexa Healthcare Skills. https://developer.amazon.com/blogs/alexa/post/ff33dbc7-6cf5-4db8-b203-99144a251a21/introducing-new-alexa-healthcare-skills (14 April 2019).

95. Food and Drug Administration. Mobile Medical Applications. 2018. https://www.fda.gov/MedicalDevices/DigitalHealth/MobileMedicalApplications/ucm255978.htm (19 Dec 2018).

96. Benjamin EJ, Munter P, Alonso A, et al. Heart disease and stroke statistics – 2019 update: a report from the American Heart Association. Circulation. 2019;139(10):e56-e528.

97. Kochanek KD, Murphy SL, Xu J, et al. Deaths: final Data for 2017. National Vital Statistics Reports. 2019;68(9).

98. De Oliveira-Filho AD, Costa FA, Neves SJ, et al. Pseudoresistant hypertension due to poor medication adherence. International Journal of Cardiology. 2014;172(2):e309-10.

99. Bailey SC, Oramasionwu CU, Wolf MS. Rethinking adherence: a health literacy–informed model of medication self-management. Journal of Health Communication. 2013;18(Suppl 1):20-30.

100. Nwankwo T, Yoon SS, Burt V, et al. Hypertension among adults in the United States: National Health and Nutrition Examination Survey, 2011-2012. NCHS Data Brief. 2013(133):1-8.

101. Dale LP, Dobson R, Whittaker R, et al. The effectiveness of mobile-health behaviour change interventions for cardiovascular disease self-management: a systematic review. Eur J Prev Cardiol. 2016;23(8):801-17.

102. Dayer LE, Shilling R, Van Valkenburg M, et al. Assessing the medication adherence app marketplace from the health professional and consumer vantage points. JMIR Mhealth and Uhealth. 2017;5(4):e45.

103. Ahmed I, Ahmad NS, Ali S, et al. Medication adherence apps: review and content analysis. JMIR Mhealth and Uhealth. 2018;6(3):e62.

104. Beleigoli AM, Andrade AQ, Cancado AG, et al. Web-based digital health interventions for weight loss and lifestyle habit changes in overweight and obese adults: systematic review and meta-analysis. J Med Internet Res. 2019;21(1):e298.

105. Khan N, Marvel FA, Wang J, et al. Digital health technologies to promote lifestyle change and adherence. Curr Treat Options Cardiovasc Med. 2017;19(8):60.106.

106. Parati G, Pellegrini D, Torlasco C. How digital health can be applied for preventing and managing hypertension. Current Hypertension Reports. 2019;21(5):40.

107. Chaix B, Bibault J, Pienkowski A, et al. When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot. JMIR Cancer. 2019;5(1):e12856.

108. Schachner T, Keller R, Wangenheim F. Artificial intelligence-based conversational agents for chronic conditions: systematic literature review. J Med Internet Res. 2020;22(9):e20701.

109. Persell SD, Peprah YA, Lipiszko D, et al. Effect of home blood pressure monitoring via a smartphone hypertension coaching application or tracking application on adults with uncontrolled hypertension a randomized clinical trial. JAMA Netw Open. 2020;3(3):e200255.

110. Migneault JP, Dedier JJ, Wright JA, et al. A culturally adapted telecommunication system to improve physical activity, diet quality, and medication adherence among hypertensive African-Americans: a randomized controlled trial. Annals of Behavioral Medicine. 2012;43(1):62-73

111. Norman DA, Draper SW. User Centered System Design: New Perspectives on Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum Associates Inc 1986.

112. Mummah SA, Robinson TN, King AC, et al. IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res. 2016;18(12):e317.

113. Mummah SA, King AC, Gardner CD, et al. Iterative development of Vegethon: a theory-based mobile app intervention to increase vegetable consumption. Int J Behav Nutr Phys Act. 2016.13(90).

114. Burchert S, Alkneme MS, Bird M, et al. User-centered app adaptation of a low-intensity e-mental health intervention for Syrian refugees. Front Psychiatry. 2018;9:663.

115. Noordman J, Driesenaar JA, van Bruinessen IR, et al. ListeningTime; participatory development of a web-based preparatory communication tool for elderly cancer patients and their healthcare providers. Internet Interv. 2017;9:51-56.

116. NIH Office of Behavioral and Social Sciences. Best practices for mixed methods research in the health sciences (2nd ed). Bethesda: National Institutes of Health. 2018.

117. Arcaya MC, Figueroa JF. Emerging trends could exacerbate health inequities in the United States. Health Affairs (Millwood). 2017;36(6):992-998.

118. Pew Research Center. Demographics of internet and home broadband usage in the United States. 2019. https://www.pewresearch.org/internet/fact-sheet/internet-broadband/ (12 April 2020).

119. Guest G, Bunce A, Johnson L. How many interviews are enough?: an experiment with data saturation and variability. Field Methods. 2006;18(1):59-82.

120. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Quality & Quantity. 2018;52(4):1893-1907.

121. Fisher WA, Fisher JD, Harman J. The information-motivation-behavioral skills model: a general social psychological approach to understanding and promoting health behavior. In J. Suls & K. A. Wallston (Eds.), Blackwell series in health psychology and behavioral medicine. Social psychological foundations of health and illness. Blackwell Publishing. 2003;82-106.

122. Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. Management Information Systems Quarterly. 2012;36(1):157-178.

123. Garavand A, Samadbeik M, Kafashi M, et al. Acceptance of health information technologies, acceptance of mobile health: a review article. J Biomed Phys Eng. 2017; 7(4):403-408.

124. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. International Journal of Medical Informatics. 2017;101:75-84.

125. van Houwelingen CT, Ettema RG, Antonietti MG, et al. Understanding older people's readiness for receiving telehealth: mixed-method study. J Med Internet Res. 2018;20(4):e123.

126. Alexander DS, Hogan SL, Jordan JM, et al. Examining whether the information–motivation–behavioral skills model predicts medication adherence for patients with a rare disease. Patient Prefer Adherence. 2017;11:75-83.

127. Jeon E, Park H. Experiences of patients with a diabetes self-care app developed based on the information-motivation-behavioral skills model: before-and-after study. JMIR Diabetes. 2019;4(2):e11590.

128. Palacio A, Garay D, Langer B, et al. Motivational interviewing improves medication adherence: a systematic review and meta-analysis. J Gen Intern Med. 2016;31(8):929-940.

129. Miller WR, Rose GS. Toward a theory of motivational interviewing. Am Psychol. 2009;64(6):527-537.

130. Miller WR, Baca LM. Two-year follow-up of bibliotherapy and therapist-directed controlled drinking training for problem drinkers. Behavior Therapy. 1983;14(3).

131. Salvo M, Cannon-Breland ML. Motivational interviewing for medication adherence. J Am Pharm Assoc. 2015;55(4):e354-361.

132. Venkatesh V, Morris MG, Davis GB, et al. User acceptance of information technology: toward a unified view. Management Information Systems Quarterly. 2003;27(3):5.

133. Lin B, Wong AM, Tseng KC. Community-based ECG monitoring system for patients with cardiovascular diseases. J Med Syst. 2016;40(4):80.

134. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Family Medicine. 2004;36(8):588-94.

135. Patient-Reported Outcomes Measurement Information System. Self-efficacy for managing medications and treatments – short form 8a. 2016. http://www.healthmeasures.net/administrator/components/com_instruments/uploads/PROMIS%20SF%20v1.0%20-%20Self-Effic-ManagMeds%208a_8-5-2016.pdf (5 August 2019).

136. Matza LS, Park J, Coyne KS, et al. Derivation and validation of the ASK-12 adherence barrier survey. Annals of Pharmacotherapy. 2009;43(10):1621-30.

137. University of Arkansas for Medical Sciences Center for Health Literacy. Patient health literacy measures. 2017. https://afmc.org/wp-content/uploads/2017/01/Literacy-Tools-UAMS-CHL-DHS-2017.pdf (5 August 2019).

138. Pact Care BV. Florence – Your health assistant. 2019. https://florence.chat/ (1 August 2019).

139. Pact Care BV. Florence. 2019. https://www.youtube.com/watch?time_continue=60&v=BtqJaHv53g0 (1 August 2019).

140. QSR International. NVivo. 2020. https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home (9 September 2019).

141. Guest G, MacQueen KM, Namey EE. Applied Thematic Analysis. Thousand Oaks, CA: Sage Publications, Inc 2012.

142. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37-46.

143. Cameron G, Cameron D, Megaw G, et al. Assessing the usability of a chatbot for mental health care. 5th International Conference on Internet Science 2018;121-132.

144. Kocaballi AB, Berkovsky S, Quiroz JC, et al. The personalization of conversational agents in health care: systematic review. J Med Internet Res. 2019;21(11):e15360.

145. Palanica A, Flaschner P, Thommandram A, et al. Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. J Med Internet Res. 2019;21(4):e12887.

146. Sara JDS, Maor E, Borlaug B, et al. Non-invasive vocal biomarker is associated with pulmonary hypertension. PLoS One. 2020.15(4):e0231441.

147. Dagum P. Digital biomarkers of cognitive function. NPJ Digit Med. 2018;1(10).

148. Office of the National Coordinator for Health Information Technology. Conceptualizing a data infrastructure for the capture, use, and sharing of patient-generated health data in care delivery and research through 2024. 2018. https://www.healthit.gov/sites/default/files/onc_pghd_final_white_paper.pdf (1 August 2019).

149. Abdolkhani R, Gray K, Borda A, et al. Patient-generated health data management and quality challenges in remote patient monitoring. JAMIA Open. 2019;2(4):471-478.

150. National Committee for Quality Assurance. Controlling high blood pressure. 2018. https://www.ncqa.org/hedis/measures/controlling-high-blood-pressure (1 August 2019).

151. Centers for Medicare and Medicaid Services. Medicare 2020 Part C & D Star Ratings Technical Notes. 2019. https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/Downloads/Star-Ratings-Technical-Notes-Oct-10-2019.pdf (1 October 2020).

152. Baptista S, Wadley G, Bird, D, et al. Acceptability of an embodied conversational agent for type 2 diabetes self-management education and support via a smartphone app: mixed methods study. JMIR mHealth uHealth. 2020;8(7):e17038.

153. Carayon P, Hoonakker P. Human factors and usability for health information technology: old and new challenges. Yearb Med Inform. 2019;28(1):71-77.

154. Yen P, Bakken S. Review of health information technology usability study methodologies. J Am Med Inform Assoc. 2012;19(3):413-422.

155. Bickmore TW, Pfeifer LM, Byron D, et al. Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. Journal of Health Communication. 2010;15 Suppl 2:197-210.

156. Bickmore TW, Caruso L, Clough-Gorr K. Acceptance and usability of a relational agent interface by urban older adults. CHI '05 Extended Abstracts on Human Factors in Computing Systems 2005:1212-1215.

157. Hauser-Ulrich S, Kunzli H, Meier-Peterhans D, et al. A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. JMIR Mhealth Uhealth. 2020;8(4):e15806.

158. Abd-Alrazaq A, Safi Z, Alajlani M, et al. Technical metrics used to evaluate health care chatbots: scoping review. J Med Internet Res. 2020;22(6):e18301.

159. Griffin AC, Xing Z, Mikles SP, et al. Information needs and perceptions of chatbots for hypertension medication self-management: a mixed methods study 2021. Journal of the American Medical Informatics Open. (In Press).

160. Google Cloud Dialogflow. API Interactions. 2020. https://cloud.google.com/dialogflow/es/docs/api-overview (9 Dec 2020).

161. World Wide Web Consortium. Web Content Accessibility Guidelines Overview. 2020. https://www.w3.org/WAI/standards-guidelines/wcag/ (11 Nov 2020).

162. U.S. National Library of Medicine. Pillbox. 2019. https://pillbox.nlm.nih.gov/ (1 Nov 2019).

163. U.S. National Library of Medicine. MedlinePlus. 2019. https://medlineplus.gov/ (1 Nov 2019).

164. Google Cloud Dialogflow. Intent Matching. 2020. https://cloud.google.com/dialogflow/es/docs/intents-matching#algo (1 Dec 2020).

165. Georgsson M, Staggers N. Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. J Am Med Inform Assoc. 2016;23(1):5-11.

166. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput. 2003;35(3):379-383.

167. Lewis C. Using the "thinking aloud" method in cognitive interface design. IBM Watson Research Center, NY. 1982.

168. Brooke J. SUS - A quick and dirty usability scale. Usability evaluation in the industry. London, England. Taylor & Francis. 1996:189-194.

169. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. Journal of Usability Studies. 2009;4(3):114-123.

170. Brooke J. SUS: a retrospective. Journal of Usability Studies. 2013;8(2):29-40.

171. Shelton CP, Koopman P, Nace W. A framework for scalable analysis and design of system-wide graceful degradation in distributed embedded systems. Proceedings of the Eighth International Workshop on Object-Oriented Real-Time Dependable Systems 2003;156-163.

172. Budiu R. The user experience of chabots. 2018. https://www.nngroup.com/articles/chatbots/ (1 Jan 2021).

173. Ragni M, Rudenko A, Kuhnert B, et al. Errare humanum est: erroneous robots in human-robot interaction. In Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). 2016;501-506.

174. Toader D, Boca G, Toader R, et al. The effect of social presence and chatbot errors on trust. Sustainability. 2020;12(1):256.

175. Ceaparu I, Lazar J, Bessiere K, et al. Determining causes and severity of end-user frustration. International and Journal of Human-Computer Interaction. 2020;333-356.

176. Cassell J, Sullivan J, Preost S, et al. Embodied conversational agents. MIT Press, Cambridge, UK 2000.

177. Bickmore TW, Pfeifer LM, Byron D, et al. Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. J Health Commun. 2010;15 Suppl 2:197-210.

178. Bickmore TW, Caruso L, Clough-Gorr K, et al. 'It's just like you talk to a friend' relational agents for older adults. Interacting with Computers. 2005;17(6):711-735.

179. Preininger AM, South B, Heiland J, et al. Artificial intelligence-based conversational agent to support medication prescribing. JAMIA Open. 2020;3(2):225-232.

180. Zand A, Sharma A, Stokes Z, et al. An exploration into the use of a chatbot for patients with inflammatory bowel diseases: retrospective cohort study. J Med Internet Res. 2020;22(5):e15589.

181. Staples S. The discourse of nurse-patient interactions: contrasting the communicative styles of U.S. and international nurses. Studies in Corpus Linguistics. 2015;72.

182. Thomas J, Short M. Using Corpora for Language Research. London, UK. Longman Pub Group 1996.

183. Adolphs S, Brown B, Carter R, et al. Applying corpus linguistics in a health care context. Journal of Applied Linguistics. 2004;1(1):9-28.

184. Microsoft Azure. Health Bot Overview. 2020. https://docs.microsoft.com/en-us/azure/health-bot/ (1 Jul 2020).

185. Mehta N, Petersen K, To W, et al. AWS for industries: building clinically-validated conversational agents to address novel coronavirus. 2020. https://aws.amazon.com/blogs/industries/building-clinically-validated-conversational-agents-to-address-novel-coronavirus/ (1 Jan 2021).

186. Bilski B, Brzozowski B, Mazur-Bialy A, et al. The role of physical exercise in inflammatory bowel disease. BioMed Research International. 2014.

187. Subramanian R, Triadafilopoulos G. Care of inflammatory bowel disease patients in remission. Gastroenterol Rep (Oxf). 2016;4(4):261-271.

188. Kochar B, Martin CF, Kappelman MD, et al. Evaluation of gastrointestinal patient reported outcomes measurement information system (GI-PROMIS) symptom scales in subjects with inflammatory bowel diseases. Am J Gastroenterol. 2018;113(1):72-79.

189. Kappelman MD, Long MD, Martin C, et al. Evaluation of the patient reported outcomes measurement information system in a large cohort of patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol. 2014;12(8):1315-23.

190. Durstine JL, Gordon B, Wang Z, et al. Chronic disease and the link to physical activity. Journal of Sport and Health Science. 2013;2(1):3-11.

191. Gatt K, Schembri J, Katsanos KH, et al. Inflammatory bowel disease and physical activity: a study on the impact of diagnosis on the level of exercise amongst patients with IBD. J Crohns Colitis. 2019;13(6):686-692.

192. Bilski J, Mazur-Bailey A, Brzozowski B, et al. Can exercise affect the course of inflammatory bowel disease? Experimental and clinical evidence. Pharmacol Rep. 2016;68(4):827-36.

193. Ng V, Millard W, Lebrun C, et al. Low-intensity exercise improves quality of life in patients with Crohn's disease. Clin J Sport Med. 2007;17(5):384-8.

194. Klare P, Nigg J, Nold J, et al. The impact of a ten-week physical exercise program on health-related quality of life in patients with inflammatory bowel disease: a prospective randomized controlled trial. Digestion. 2015;91(3):239-47.

195. Eckert KG, Abbasi-Neureither I, Koppel M, et al. Structured physical activity interventions as a complementary therapy for patients with inflammatory bowel disease – a scoping review and practical implications. BMC Gastroenterol. 2019;19:115.

196. Ho GWK. Lower gastrointestinal distress in endurance athletes. Curr Sports Med Rep. 2009;8(2):85-91.

197. Pricewaterhouse Coopers. The wearable life 2.0: connected living in a wearable world. 2018. https://www.pwc.com/ee/et/publications/pub/pwc-cis-wearables.pdf (10 Jan 2020).

198. Pew Research Center. About one-in-five Americans use a smart watch or fitness tracker. 2020. https://www.pewresearch.org/fact-tank/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/ (10 March 2020).

199. Marschollek M. A semi-quantitative method to denote generic physical activity phenotypes from long-term accelerometer data – the ATLAS index. PLoS One. 2013;8(5):e63522.

200. Lee I, Sesso HD, Oguma Y, et al. The "weekend warrior" and risk of mortality. Am J Epidemiol. 2004;160(7):636-641.

201. Metzger JS, Catellier DJ, Evenson KR, et al. Patterns of objectively measured physical activity in the United States. Med Sci Sports Exerc. 2008;40(4):630-638.

202. McConnell MV, Shcherbina A, Pavlovic A, et al. Feasibility of obtaining measures of lifestyle from a smartphone app: the MyHeart Counts cardiovascular health study. JAMA Cardiol. 2017;2(1):67-76.

203. U.S. Department of Health and Human Services. Physical activity guidelines for americans. 2nd edition. 2018. https://health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf#page=56. (15 Feb 2021).

204. Sossenheimer PH, Yvellez OV, Andersen M, et al. Wearable devices can predict disease activity in inflammatory bowel disease patients. Journal of Crohn's and Colitis. 2019;13(Supplement 1):S404.

205. Wiestler M, Kockelmann F, Kuck M, et al. Quality of life is associated with wearable-based physical activity in patients with inflammatory bowel disease: a prospective, observational study. Clin Transl Gastroenterol. 201910(11):e00094.

206. Long MD, Kappelman MD, Martin CF, et al. Development of an internet-based cohort of patients with inflammatory bowel diseases (CCFA Partners): methodology and initial results. Inflamm Bowel Dis. 2012;18(11):2099-106.

207. Thia K, Faubion WA, Loftus EV, et al. Short CDAI: Development and validation of a shortened and simplified Crohn's disease activity index. Inflamm Bowel Dis. 2010;17(1):105-111.

208. Jowett SL, Seal CJ, Phillips E, et al. Defining relapse of ulcerative colitis using a symptom-based activity index. Scand J Gastroenterol. 2003;38(2):164-71.

209. Higgins PDR, Schwartz M, Mapili J, et al. Patient defined dichotomous end points for remission and clinical improvement in ulcerative colitis. Gut. 2005;54(6):782-788.

210. Tukey JW. Exploratory Data Analysis. 1st ed. Pearson 1977.

211. Lloyd S. Least Squares Quantization in PCM. IEEE Transactions on Information Theory. 2006;28(2).

212. Tan P, Steinbach M, Kumar V. Introduction to data mining. Boston, MA: Pearson Addison Wesley 2005.

213. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. Journal of Machine Learning Research. 2011:2825-2830.

214. Arthur D, Vassilvitskii S. k-means++: the advantages of careful seeding. SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms 2007;1027-1035.

215. Reedy J, Wirfalt E, Flood A, et al. Comparing 3 dietary pattern methods—cluster analysis, factor analysis, and index analysis—with colorectal cancer risk: the NIH–AARP diet and health Study. Am J Epidemiol. 2010;171(4):479-87.

216. Jones PD, Kappelman MD, Martin CF, et al. Exercise decreases risk of future active disease in inflammatory bowel disease patients in remission. Inflamm Bowel Dis. 2015 21(5):1063-1071.

217. Kamp KJ, West P, Holmstrom A, et al. Systematic review of social support on psychological symptoms and self-management behaviors among adults with inflammatory bowel disease. J Nurs Scholarsh. 2019;51(4):380-389.

218. Conley S, Redeker N. A systematic review of self-management interventions for inflammatory bowel disease. J Nurs Scholarsh. 2016;48(2):118-127.

219. Stephenson A, McDonough SM, Murphy MH, et al. Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act. 2017;14(105).

220. Regueiro M, Greer JB, Szigethy E. Etiology and treatment of pain and psychosocial issues in patients with inflammatory bowel diseases. Gastroenterology. 2016152(2):430-439.

221. Regueiro MD, McAnallen SE, Greer JB, et al. The inflammatory bowel disease specialty medical home: a new model of patient-centered care. Inflamm Bowel Dis. 2016;22(8):1971-80.

222. Chung AE, Basch EM. Potential and challenges of patient-generated health data for high-quality cancer care. J Oncol Pract. 2015;11(3):195-197.

223. Li X, Dunn J, Salins D, et al., Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. PLoS Biol. 2017;15(1).

224. Radin JM, Wineinger NE, Topol EJ, et al. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. Lancet Digital Health. 2020. 2(2):e85-e93.

225. Dunn J, Runge R, Snyder M. Wearables and the medical revolution. Personalized Medicine. 2018;15(5):429-448.

226. Althoff T, Sosic R, Hicks JL, et al. Large-scale physical activity data reveal worldwide activity inequality. Nature. 2017;547(7663):336-339.

227. Nass C, Steuer J, Tauber ER. Computers are social actors. CHI ’94 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 1994:72-78.

228. Luxton DD. Ethical implications of conversational agents in global public health. Bulletin of the World Health Organization. 2020;98:285-287.

229. Zhang J, Oh YJ, Lange P, et al. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: viewpoint. JMIR. 2020;22(9).

230. Weizenbaum J. Computer power and human reason: from judgment to calculation. ed. San Francisco, CA, W.H. Freeman 1976.

231. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. Artificial Intelligence in Medicine. 2014;62(1):1-10.

232. World Economic Forum. Chatbots RESET A framework for governing responsible use of conversational AI in healthcare. 2020. https://www.weforum.org/reports/chatbots-reset-a-framework-for-governing-responsible-use-of-conversational-ai-in-healthcare (17 Jan 2021).

233. International Medical Device Regulators Forum. Software as a medical device: possible framework for risk categorization and corresponding considerations. 2014. https://www.fdanews.com/ext/resources/files/04/04-07-14-software.pdf (12 Dec 2020).

234. Franssen WMA, Frannse GHL, Spaas J, et al. Can consumer wearable activity tracker-based interventions improve physical activity and cardiometabolic health in patients with chronic diseases? A systematic review and meta-analysis of randomised controlled trials. International Journal of Behavioral Nutrition and Physical Activity. 2020;17(57).

235. HL7 FHIR. Enabling health interoperability through FHIR. 2020. http://fhir.org/ (16 Jan 2021).

236. Roca S, Sancho J, García J, et al. Microservice chatbot architecture for chronic patient support. Journal of Biomedical Informatics. 2020; 102(103305).

237. Sayeed R, Gottlieb D, Mandl KD. SMART Markers: collecting patient-generated health data as a standardized property of health information technology. NPJ Digital Medicine. 2020;3(9).

238. Li RC, Asch SM, Shah NH. Developing a delivery science for artificial intelligence in healthcare. NPJ Digital Medicine. 2020;3(107).

239. Carayon P, Hundt AS, Karsh BT, et al. Work system design for patient safety: the SEIPS model. Qual Saf Health Care. 2006;15(Suppl 1): i50-i58.

240. Omron. Wearable blood pressure monitor. 2020. https://omronhealthcare.com/products/heartguide-wearable-blood-pressure-monitor-bp8000m/ (2 Feb 2021).

241. Okunrintemi V, Spatz ES, Capua PD, et al. Patient–provider communication and health outcomes among individuals with atherosclerotic cardiovascular disease in the United States. Circulation Cardiovascular Quality and Outcomes. 2017;10(4):e003635.

242. Torous J, Myrick KJ, Rauseo-Ricupero N, et al. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health. 2020;7(3).

243. Judson TJ, Odisho AY, Neinstein AB, et al. Rapid design and implementation of an integrated patient self-triage and self-scheduling tool for COVID-19. J Am Med Inform Assoc. 2020;27(6):860-866.

244. John Hopkins Medicine. Coronavirus (COVID-19) self-checker. 2020. https://www.hopkinsmedicine.org/coronavirus/covid-19-self-checker.html (18 April 2020).

245. St. Joe's Health System. COVID-19 symptom checker. 2020. https://www.stjoeshealth.org/health-and-wellness/coronavirus (18 April 2020).

246. Emory Healthcare. COVID-19 symptom checker. 2020. https://www.emoryhealthcare.org/covid/symptom-checker.html (19 April 2020).

247. Cleveland Clinic. Find out your COVID-19 risk. 2020. https://my.clevelandclinic.org/landing/preparing-for-coronavirus (19 April 2020).

248. World Health Organization. WHO launches a chatbot on Facebook Messenger to combat COVID-19 misinformation. 2020. https://www.who.int/news-room/feature-stories/detail/who-launches-a-chatbot-powered-facebook-messenger-to-combat-covid-19-misinformation (17 April 2020).

249. Centers for Disease Control and Prevention. Symptoms of coronavirus. https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html (17 April 2020).

250. Siwicki B. Northwell, UCSF, UNC using chatbot and related tech to manage COVID-19 patients. 2020. https://www.healthcareitnews.com/news/northwell-ucsf-unc-using-chatbot-and-related-tech-manage-covid-19-patients (25 June 2020).

251. Pfefferbaum B, North CS. Mental health and the covid-19 pandemic. N Engl J Med. 2020;383:510-512.

252. Muoio D. Apple's Siri now walks worried users through their COVID-19 symptoms. 2020. https://www.mobihealthnews.com/news/apples-siri-now-walks-worried-users-through-their-covid-19-symptoms (20 April 2020).

253. Amazon. Alexa and Amazon devices COVID-19 resources. 2020. https://blog.aboutamazon.com/devices/alexa-and-amazon-devices-covid-19-resources?ots=1&tag=curbedcom06-20&linkCode=w50 (17 June 2020).

254. Babylon Health. 2020. https://www.babylonhealth.com/ (2 Feb 2020).