Managing Chronic Conditions with a Smartphone-based Conversational Virtual Agent

Timothy W. Bickmore
Northeastern University
Boston, MA, USA
bickmore@ccs.neu.edu

Everlyne Kimani
Northeastern University
Boston, MA, USA
kimani15@ccs.neu.edu

Ha Trinh
Northeastern University
Boston, MA, USA
hatrinh@ccs.neu.edu

Alexandra Pusateri
University of Pittsburgh Med Center
Pittsburgh, PA, USA
pusateriam@upmc.edu

Michael K. Paasche-Orlow
Boston Medical Center
Boston, MA, USA
michael.paasche-orlow@bmc.org

Jared W. Magnani
University of Pittsburgh Med Center
Pittsburgh, PA, USA
magnanijw@upmc.edu

ABSTRACT

When deployed on mobile devices, virtual agents have the potential to deliver advice regarding medical conditions, as well as provide a ubiquitous channel for health education and behavior change for a variety of chronic health conditions. We describe design guidelines for mobile agent dialogues to support chronic disease management, a general-purpose smartphone-based architecture for a conversational virtual agent that simulates face-to-face health counseling conversations with patients, and an initial agent implementation that provides counseling to patients with the chronic heart condition atrial fibrillation in conjunction with a mobile heart rhythm monitor that is attached to the back of the phone. Preliminary results from a randomized trial with 120 patients with atrial fibrillation indicate that the agent results in significant improvements in self-reported quality of life relative to a standard of care control group.

CCS CONCEPTS

• Human-Computer Interaction → Usability and acceptability, Mobile applications, Conversational agents

KEYWORDS

Health, Human-centered computing, Mobile Applications, Natural language interfaces

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

ACM Reference Format:


1 INTRODUCTION

Chronic conditions, such as diabetes, hypertension, and atrial fibrillation, affect about half of all adults in the US, and account for 7 of the top 10 causes of death. These conditions require challenging self-care management, spanning medication adherence, lifestyle modifications, and vigilant symptom monitoring in addition to regular visits with health-care providers to prevent disease progression. However, adherence to these regimens is generally poor, especially for the 36% of adults with limited health literacy who struggle to understand and act on instructions from their providers.

Smartphones provide an ideal platform for systems designed to help individuals manage their chronic conditions, given their computational power and connectivity, their “anywhere, anytime” availability, and especially when coupled with mobile sensors or monitors. However, many individuals struggle with complex user interfaces, and while the omnipresence of these devices removes some adherence barriers, most applications still fail to address many other barriers, such as lack of motivation or self-efficacy.

Over the last decade, a series of automated health-care interventions that use conversational agents in the role of health coaches have been developed to provide health education and promote behavior change. Conversational agents provide an accessible interface for users with limited health or computer literacy, and provide additional channels of communication that are especially effective for building trust and therapeutic alliance...
to promote adherence. These agents have also been deployed on mobile devices with integrated sensors, to provide anywhere, anytime health counseling.

We describe recent work in developing a general architecture for a smartphone-based virtual agent for chronic disease management, and an initial application in promoting self-care for individuals with atrial fibrillation. Atrial fibrillation (AF) is an irregular heartbeat that increases the risk of stroke 3- to 5-fold, and doubles the risk of death if untreated. AF management includes medication adherence and vigilant symptom monitoring, and is unusual in that patients can have trouble interpreting their symptoms: palpitations can be present without AF, and AF can be present without palpitations. The recent advent of mobile heart rhythm monitors enables individuals to correlate symptoms with heart rate and rhythm. However, this requires patients to take regular readings as part of their self-care regimen [1].

Pilot studies of earlier versions this system were reported in [2] and [3]. This paper goes beyond earlier work in reporting the design principles for virtual agent coaches for chronic disease management, and the design and architecture of the AF system, extending the intervention from one week to 30 days and reporting results from both a preliminary study and a randomized pilot involving 120 patients with atrial fibrillation.

2 RELATED WORK

2.1 Mobile Interventions for Chronic Disease Management

Several technology-based interventions for chronic disease management have been explored in the HCI community over the last 20 years, with most work focusing on diabetes management. For example, Mamykina, et al, describe MAHI, a system for enabling individuals with diabetes to reflect on their diet and blood glucose levels, finding that it was effective in promoting diet change, blood glucose monitoring, and exercise frequency [4]. Since then, dozens of commercial smartphone apps have become available with similar functionality. These have mostly focused on providing tools to enable users to collect and reflect on manually entered biometric, behavioral, and self-report data. Unfortunately, manual entry often represents a significant barrier [5]. In addition, a study of six commercial “collect and reflect” smartphone apps for diabetes management were evaluated and found to be too complex for use on mobile displays [6, 7].

Several mobile devices, including smartphones, watches, and sensors that interoperate with these, have now been developed to detect atrial fibrillation. Ding, et al, conducted a study to assess impressions of these devices by older patients, finding that most patients (78%) found the devices tested acceptable and easier to use than traditional cardiovascular monitors [8].

2.2 Mobile Virtual Agents

Bickmore, et al, conducted a design study demonstrating that an animated virtual agent on a handheld device was more effective at building trust with users than equivalent static agent images or text-only interfaces [9]. Kang, et al, describe similar studies investigating user reactions to an animated virtual agent on a smartphone compared to a static agent image or no image, and find that users have longer conversations with the animated agent [10] and rated agents with more anthropomorphic appearances higher on satisfaction and perceived co-presence [11].

2.3 Virtual Agent Health Interventions
A number of conversational agents have now been developed to counsel patients on health problems, in general, and chronic disease self-care, in particular, although few have been deployed on mobile devices. Bickmore, et al, investigated the use of a virtual exercise counselor agent on a PDA device with integrated pedometer [12]. Leuski, et al, describe a virtual animated agent on a smartphone that helps diagnose medical conditions, although the system described is an incomplete concept demonstration [13].

2.4 AF Interventions

Regarding AF specifically, there are very few automated interventions described in the literature to date for promoting adherence or self-management skills. The “Health Buddies” intervention used behavioral contracting and accountability between patients with AF and their grandchildren to motivate adherence [14]. A feasibility study with 15 patients demonstrated acceptance, usability, and (non-significant) knowledge gain by users. However, only 13% of eligible patients agreed to participate, with lack of interest in the concept or discomfort using conventional technology being key reasons for declining.

3 MOBILE AGENT DESIGN

The design of our agent for chronic disease management was derived from several guidelines and design principles. Our use of a virtual agent was motivated by studies demonstrating that agent-based interfaces are significantly more usable by individuals with low health and computer literacy compared to more conventional interfaces [15].

3.1 Requirements for Chronic Disease Management

There are several care models for chronic disease management [16]. The most-cited framework is the Chronic Care Model, which describes six elements that are needed to provide proactive, patient-centered, evidence-based care [17]. Of these, three are particularly relevant to automated interventions. First, systems should provide patient tracking, be able to alert healthcare providers when needed, and should facilitate the exchange of information between providers and patients, requiring that mobile interventions need to be able to communicate with clinical information systems. Second, care guidelines consistent with scientific evidence and patient preferences should be shared with patients to encourage participation in medical care. Finally, patients should be motivated and supported in managing their health, including goal setting, action planning, skill development, and various lifestyle behaviors.

3.2 Requirements for Mobile Health Interventions

Based on evaluations of commercial smartphone apps for diabetes management, we know that providing visualizations of large amounts of data for reflection can be counterproductive on small displays [6, 7], indicating that the interface should be kept as simple as possible. E-health literacy and privacy concerns have also been shown to influence adoption of health apps. Therefore, in addition to simplified interfaces and interaction modalities for people with varying e-health literacy, the app and its interaction modality should mitigate privacy concerns. Other guidelines for the design of mobile interfaces in general indicate that interactions should be user-directed and interruptible given that users may be interweaving smartphone and non-smartphone tasks [18].

3.3 Design Guidelines for Mobile Virtual Agent Health Counselors

Based on the above requirements, we derived a set of general guidelines for the design of mobile conversational coaches for chronic condition management. Support short interactions. Mobile interactions are often of short duration, due to the high frequency of interruption and distraction from other activities. Allow for interruption. As interruption is a common issue in mobile interactions, agent dialogue should incorporate mechanisms to maintain continuity in the conversation when interruptions occur. Use pre-suggested user responses. Inaccuracies in Speech Recognition and Natural Language Understanding can lead to dangerous consequences in safety-critical applications such as healthcare. Thus, user input should be fully constrained (e.g., via multiple-choice menus). Maintain persistence across sessions. To support longitudinal health counseling and demonstrate continuity in the therapeutic relationship with the agent, the agent should maintain a memory of past interactions with the user and dynamically tailor the current conversation accordingly.

3.4 System Architecture

Our smartphone-based virtual agent framework consists of four components: an embodied conversational agent, a custom dialogue engine, a platform-specific text-to-speech engine, and a database store.

Our conversational agent is a 3D-animated humanoid character that communicates with the user through synthetic speech and synchronized nonverbal conversational behavior. The character’s appearance can be tailored for specific user populations using a commercial character design tool. To emulate face-to-face conversations, the animated agent can display a variety of nonverbal behavior, including: head nods; eyebrow raises, directional gazes for signaling turn-taking; a range of hand gestures (e.g., beat gestures for emphasis, contrastive gestures for comparisons), and posture shifts to mark topic boundaries, all generated using the BEAT text-to-embodied speech system [19].

Human-agent dialogues are authored using a hierarchical transition network-based scripting language, augmented with template-based text generation. User responses are primarily in
ATRIAL FIBRILLATION VIRTUAL AGENT

Our first mobile chronic disease management virtual agent is designed to help individuals with AF. The agent application is installed on a smartphone, and designed to be used in conjunction with a mobile heart rhythm monitor attached to the phone (Fig. 1).

4.1 Mobile Heart Rhythm Monitor

Frequent monitoring of heart rhythm is crucial for people with AF. Patients with AF often have to schedule multiple visits with a cardiologist to get their heart rhythm recorded and assessed, but recent development of mobile heart rhythm monitors is helping patients monitor their heart rhythm with their phone. In our intervention we provide all participants with an AliveCor Kardia mobile heart rhythm monitor (Fig. 1, Right). This monitor has been validated for use as an AF screening tool in clinical research studies.

The AliveCor monitor is attached to the back of a smartphone and data is transmitted via Bluetooth connection to the phone via heart rhythm recording AliveCor app which provides an automated analysis. The AliveCor app also allows users to include notes about any symptoms they might have at the time of a recording as well as share their heart rhythm report with their doctors via email.

4.2 Virtual Agent Dialogue Topics

We designed our virtual agent dialogue according to the design principles in section 3.3 and to be accessible to patients who have been newly diagnosed with AF and who have a wide range of literacy levels. The dialogue covers the following topics:

AF education, including: overview of AF and causes and consequences of AF; AF treatments and their indications; an overview of self-care for chronic illness; self-monitoring skills; and the importance of self-care management.

Kardia heart rhythm monitor education, including: how and when to use the monitor; how to distinguish between a regular heart rhythm and a heart rhythm in AF; and, how to interpret and act on results. The agent also promotes adherence to daily heart rhythm monitoring by reminding patients of the importance of taking regular readings and helping resolve barriers to regular use.

Symptom education and reporting, including: common symptoms such as palpitations, shortness of breath and chest pain, as well as an assessment of frequency and severity of these symptoms; association of symptoms with AF; treatment of symptoms; and actions to take.

**Medications:** Most AF patients are prescribed an anticoagulant medication which prevents blood clots and strokes. This topic includes: rationale for anticoagulation medications in AF; the importance of taking AF medication as prescribed, and common barriers to medication adherence such as busyness, forgetfulness and poor access to pharmacies, along with recommendations on how to deal with these challenges.

**Medication side-effects,** including: common side effects of anticoagulation medications; and treatment of side-effects.

**Emergency conditions,** including: recognition of emergencies and actions to take.

**Patient activation,** motivating the user to take charge of their own health, and preparing for medical encounters.

During the first week of the intervention, the virtual agent prioritizes education and common symptom reporting. After the first week of interaction, the content covered and flow of dialogue is primarily user-directed, although the virtual agent suggests a topic at the beginning of each interaction if the topic has not already been covered. Patients also have the option to review previously covered topics.

The agent periodically asks users about their quality of life, assesses changes in symptom frequency and severity, changes in lifestyle after their AF diagnosis, and their level of knowledge about AF, with results from these assessments saved for review. Although our current implementation does not support real-time alerting of healthcare providers, the database synchronization architecture can be easily extended to support this.

5 PILOT STUDY

We conducted an initial 30-day quasi-experimental demonstration study to assess usability of our AF agent, prior to starting a clinical trial.

In addition to sociodemographics, system use was logged on central servers, and self-report measures of satisfaction were assessed at the end of the 30 days, along with a semi-structured interview. Thirty-one participants with AF were recruited from outpatient clinics at the University of Pittsburgh Medical Center (UPMC). Participants were 68 (SD 11) years old, 39% female, 94% white, and 23% had inadequate health literacy based on a standardized test (S-TOFHLA [20]). Participants were provided with iPhones with the Virtual Agent app pre-installed (Fig. 1), along with integrated Kardia monitors, and provided with training on their use. They were contacted at days 7, 14, and 21 to assist with any problems they may have been having. At 30 days they returned for outcome assessment and interviews.

Study participants used the coach agent from 3 to 30 days (median 17). The number of log-ins to the agent app ranged from 4 to 43 (median of 20). The average length of interactions ranged from 1.3 minutes to 7.75 minutes (mean 2.42 minutes). The median number of Kardia uses was 29 (ranging 5 to 30). Participants using the Kardia were in AF an average of 14.3 (SD
Most participants (77%) reported some symptoms to the agent during the month. Satisfaction with the agent was high, with 13 of 31 participants reporting either 6 or 7 on a 7-point scale (median = 5, range 1-7, with 1 = “very unsatisfied” to 7 = “very satisfied”). Participants with inadequate health literacy reported a significantly higher desire to continue using the agent at the end of the 30 days, Mann-Whitney U=41.5, p<.05.

**Participant Reactions.** We conducted a qualitative analysis of the interview transcripts using thematic analysis techniques. Our analysis started with open coding, followed by clustering of relevant codes into common concepts and themes. Five themes emerged characterizing how participants felt about using the conversational virtual agent in conjunction with the mobile heart rhythm monitor to manage their AF.

**Perception of virtual agent.** To understand how participants felt about the virtual agent we conducted a sentiment analysis on interview responses in which participants expressed opinions about the agent. Using a word-based scoring system (in the NVivo qualitative analysis tool), we classified utterances as positive, negative, or mixed. Over 89.7% of utterances expressed positive opinions about the agent, while only 4.4% were negative, and 5.9% were mixed.

Most participants reported that the agent was friendly and appropriate: “She was friendly and polite, so I had no problem doing my part in that relationship” [P6]. The agent was also found to be therapeutic and comforting: “I live alone so there’s no one asking me how I’m feeling, so there are sometimes I can go days without connecting with anyone. But, I liked that you know... I could have some sort of comfort from it” [P6]. Most participants also felt that the agent provided a sense of reassurance: “Very reassuring. It was like having a doctor with me” [P3]. Regarding the virtual agent dialogue, most users indicated that the information provided was informative and easily accessible: “She’s so knowledgeable. With my mind, I’m not into medicine as a professional would be but whoever thought of it was very thorough. It’s like an answer for everything I might want to tell a health professional if they were living and breathing right next to me. It’s tremendous” [P11]. They also appreciated the personalized dialogue: “That someone was calling you by name and ‘well is there anything else I can help you with today?’ It was pretty cool to work with” [P12]. Although participants felt that the virtual agent was informative, they pointed out that it was repetitive. They suggested including additional content related to nutrition, and emotional and social aspects of living with AF.

Most participants found the virtual agent interaction modality to be easy to use: “You made it easy and I appreciated that. It was a piece of cake. User friendly so I wouldn’t change a thing” [P6]. They also stated that the agent’s voice was clear and understandable: “… I thought the avatar was well designed. I understood her…” [P1]. Users liked having the agent on a smartphone; “I liked talking on the phone, that was good” [P11].

**Interaction patterns.** Users reported similar habits when using the virtual agent and the heart rhythm monitor. All users recounted that they interacted with the virtual agent and used the heart rhythm monitor almost every day at home, integrating them into their daily routines: “Once a day ... As soon as I get up in the morning, I go to the chair and use it before I get my coffee” [P10]. Interestingly, we found that only two users reported using the system outside their homes and only one user interacted with it in a public setting, despite the system’s mobility: “Oh yeah, well people were amazed that a telephone did that. Especially at Church they would see me toying with it and taking a reading and they said oh what that is? I said ‘watch this’” [P21].

**Interaction motivation.** Our analysis revealed two main reasons that participants interacted with the virtual agent. Some reported that they used and would continue to use the virtual agent for its pedagogical content. For example, P4 stressed: “I used it to really - I wanted to learn, I wanted to know what was going on and I didn’t do it out of any altruistic reason”. When asked why he would have continued interacting with the agent, P4 said, “If I could increase my knowledge, if it had more content”. Other participants reported that they mainly interacted with the virtual agent to discuss their symptoms. Additionally, they reported that they would have continued interacting with the virtual agent if or when they had symptoms: “probably if I started to have some issues but other than that I’m pretty good to go with my knowledge of the AFib I’m good to go.” [P16].

### 6 RANDOMIZED TRIAL

Following the successful demonstration, we have recently completed a 30-day between-subjects trial to evaluate the usability and longitudinal use of the AF agent, compared to standard care, for patients with AF. Here we report preliminary results. The study was approved by the UPMC IRB.

In addition to the pilot study measures (Section 5), we included a standard measure of quality of life: Treatment Satisfaction from the AF-specific Atrial Fibrillation Effect on QualiTy of life, or AFEQT [21].

One hundred and twenty (120) participants with AF were recruited from outpatient clinics at UPMC (61 randomized to the agent intervention). Participants were 72.1 (SD 9.10) years old, and 51.7% female. Participants were well-educated, with fewer than 10% with inadequate health literacy based on the Short Test of Functional Health Literacy in Adults measure (S-TOFHLA).

All participants were administered baseline assessments at the start of the study and outcome assessments and interviews after 30 days. Participants in the intervention group were provided with iPhones with the Virtual Agent app pre-installed (Fig. 1), along with integrated Kardia monitors, and provided with training on their use. They were contacted by phone at days 7, 14, and 21 to assist with any problems they may have been having. Participants in the control group received standard care.

**Results.** Intervention group participants had 1 to 45 complete conversations with the coach agent (mean 16.2, SD 11.2) over the 30 days of the intervention. Participants logged a total of 2.1 to 179.5 (mean 41.9, SD 30.1) minutes of interaction, averaging 2.1
(SD 1.0) minutes per conversation. During their interactions, participants reported 0 to 14 (mean 1.3, SD, 2.3) symptoms to the agent.

Participants in the intervention group reported significantly higher quality of life on the AFEQT after 30 days, compared to those in the control group, 81.5 (SD 14.2) vs. 76.0 (SD 17.6), p<.05, using an ANCOVA controlling for baseline scores.

7 CONCLUSIONS

We conducted 30-day demonstration and a randomized pilot evaluation of an intervention that combined a smartphone-based virtual agent and a mobile heart rhythm monitor to help patients with AF manage their condition. Our findings highlight how users with AF interacted with our smartphone-based virtual agent. We found that study participants with regularly used the agent over the month of the study, reported symptoms, and were satisfied with the agent. Participants who interacted with the agent saw significant improvements in their self-reported quality of life relating to AF compared to those in the control group.

We conclude by discussing the degree to which the results of the usage and interaction patterns of users provided support for our design principles in Section 3.3.

Support short interactions. From the mobile coach usage results, we found that the average length of interaction was 2.4 (SD 1.3) minutes in the demonstration and 2.1 (SD 1.0) minutes in the pilot, and the dialogue topics that required more than 5 minutes to complete had more interruptions.

Allow for interruption. In our framework, we pre-define specific interruption timeouts of 15 minutes. If the dialogue is interrupted and later resumed within that time limit, the agent asks the users if they would like to continue where they left off.

Support routines and motivations. Participants had different habits when using the virtual coach. For instance, from the interviews we learned that most participants interacted with the virtual coach either in the morning or in the evening. Additionally, some participants reported that they used the coach when they were taking their medications. The mobile agent could learn these interactions patterns and prompt the user for an interaction when they are most likely to engage in a conversation about their condition.

Limitations of these studies include relatively small convenience samples, and the lack of long-term, objectively-measured health outcomes.

Given that chronic diseases affect such a large portion of the population, and the complexity of self-care management regimens—especially for patients with low health literacy—virtual agents represent an important tool for improving population health and decreasing healthcare costs.

ACKNOWLEDGMENTS

This work was supported by Grant 2015084 from the Doris Duke Charitable Foundation.

REFERENCES


