



## Research and Applications

# A chatbot for hypertension self-management support: user-centered design, development, and usability testing

Ashley C. Griffin , PhD, MSPH<sup>1,2,\*</sup>, Saif Khairat , PhD, MPH, FAMIA<sup>3,4</sup>,  
Stacy C. Bailey, PhD, MPH<sup>5</sup>, Arlene E. Chung, MD, MHA, MMCi, FAMIA<sup>6</sup>

<sup>1</sup>VA Palo Alto Health Care System, Palo Alto, CA 94025, United States, <sup>2</sup>Department of Health Policy, Stanford University School of Medicine, Stanford, CA 94305, United States, <sup>3</sup>Carolina Health Informatics Program, University of North Carolina at Chapel Hill (UNC), Chapel Hill, NC 27599, United States, <sup>4</sup>School of Nursing, UNC, Chapel Hill, NC 27599, United States, <sup>5</sup>Division of General Internal Medicine, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, United States, <sup>6</sup>Department of Biostatistics & Bioinformatics, Duke School of Medicine, Durham, NC 27710, United States

\*Corresponding author: Ashley C. Griffin, PhD, MSPH, VA Palo Alto Health Care System (152-MPD), 795 Willow Road, Menlo Park, CA 94025 (griffina@stanford.edu)

## ABSTRACT

**Objectives:** Health-related chatbots have demonstrated early promise for improving self-management behaviors but have seldomly been utilized for hypertension. This research focused on the design, development, and usability evaluation of a chatbot for hypertension self-management, called “Medicagent.”

**Materials and Methods:** A user-centered design process was used to iteratively design and develop a text-based chatbot using Google Cloud’s Dialogflow natural language understanding platform. Then, usability testing sessions were conducted among patients with hypertension. Each session was comprised of: (1) background questionnaires, (2) 10 representative tasks within Medicagent, (3) System Usability Scale (SUS) questionnaire, and (4) a brief semi-structured interview. Sessions were video and audio recorded using Zoom. Qualitative and quantitative analyses were used to assess effectiveness, efficiency, and satisfaction of the chatbot.

**Results:** Participants ( $n = 10$ ) completed nearly all tasks (98%, 98/100) and spent an average of 18 min (SD = 10 min) interacting with Medicagent. Only 11 (8.6%) utterances were not successfully mapped to an intent. Medicagent achieved a mean SUS score of 78.8/100, which demonstrated acceptable usability. Several participants had difficulties navigating the conversational interface without menu and back buttons, felt additional information would be useful for redirection when utterances were not recognized, and desired a health professional persona within the chatbot.

**Discussion:** The text-based chatbot was viewed favorably for assisting with blood pressure and medication-related tasks and had good usability.

**Conclusion:** Flexibility of interaction styles, handling unrecognized utterances gracefully, and having a credible persona were highlighted as design components that may further enrich the user experience of chatbots for hypertension self-management.

## LAY SUMMARY

Chatbots have demonstrated early potential to deliver evidence-based, interactive self-management interventions. As little research has focused on the use of chatbots for hypertension, we leveraged a user-centered design process to design, develop, and evaluate the usability of a chatbot for hypertension self-management, called “Medicagent.” Among 10 participants who completed usability testing, a mean of 18 min (SD = 10) was spent completing 10 tasks within Medicagent. Nearly all tasks (98%, 98/100) were completed. Approximately 9% (11/128) of utterances were not successfully mapped to an intent due to unrecognized spelling or formatting of dates, times, and blood pressure values. Medicagent achieved a mean System Usability Scale score of 78.8/100, which demonstrates acceptable usability. Several participants had difficulties navigating the conversational interface without menu and back buttons, felt additional information would be useful for redirection when utterances were not recognized, and desired a health professional persona to establish credibility. This study may provide insights for the design and development of chatbots to enrich the patient experience and inform components of health-related chatbots for other chronic conditions requiring self-management.

**Key words:** hypertension; self-management; chatbot; conversational agent; mobile health.

## Background and significance

Hypertension is the most common chronic condition in the United States and affects almost half of adults (116 million).<sup>1,2</sup> As a leading contributor of heart disease and stroke, patients with hypertension require ongoing self-management to prevent complications.<sup>1</sup> These self-management skills involve monitoring blood pressure, taking medications as

prescribed, working with the care team, and maintaining a healthy lifestyle.<sup>3</sup> Self-management and behavior modifications are challenging, and a number of digital health tools to support hypertension self-management exist, such as digital pill boxes, mobile health applications (apps), and more recently, chatbots.<sup>4–6</sup> Chatbots, which are dialogue systems that mimic human chat characteristics,<sup>7</sup> have emerged on

familiar platforms (Amazon Alexa, Telegram, websites, etc.) and have increasingly been adopted for health use cases.<sup>8-12</sup> One of the earliest chatbots, ELIZA, communicated like a Rogerian psychotherapist through empathic and reflective messaging in the 1960s.<sup>13</sup> With recent advances in artificial intelligence, dialogue systems have begun to move from rule-based systems to statistical data-driven systems that learn from corpus-based data.<sup>14</sup> These computational innovations, along with the growing burden of chronic disease, have ushered in a growing ecosystem of patient-facing chatbots for mobile health coaching and education.<sup>8-10</sup>

Early research studied the use of chatbots for hypertension self-management.<sup>15-19</sup> To our knowledge, only 3 randomized controlled trials have evaluated the efficacy of hypertension-related dialogue systems.<sup>15-17</sup> These systems demonstrated improvements in self-confidence in controlling blood pressure, knowledge and skills of home blood pressure monitoring, and diet quality and energy expenditure.<sup>15-17</sup> However, no differences were found in blood pressure control or treatment adherence between intervention and control groups.<sup>15-17</sup> These initial findings may suggest that the ideal design and user interactions of chatbots to support hypertension self-management are not well delineated.

Although there is limited research, self-management chatbots have been reported as acceptable and usable by most patients thus far.<sup>20-27</sup> Older adults and those with low health and computer literacy have expressed positive attitudes and found them easy to use.<sup>20-22</sup> To assess how chatbots can help users achieve their goals, 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, realism, and satisfaction.<sup>28</sup> These metrics provide methods for understanding and improving the usability of chatbots to better meet patients' needs. Usability challenges present a major barrier to health information technology adoption<sup>29</sup> and prevent users from achieving their goals with efficiency, effectiveness, and satisfaction.<sup>30</sup> Involving patients throughout design processes and ensuring technologies are usable have illustrated positive improvements in patient empowerment and safety.<sup>31</sup> In this study, we leveraged a user-centered design process to design and develop a chatbot focused on hypertension self-management. We then evaluated 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 prototype, called "Medicagent." We also examined strengths and shortcomings of the chatbot to inform optimizations in preparation for future pilot testing.

## Methods

### Chatbot design

Throughout the design process, we used components of the Integrate, Design, Assess, and Share (IDEAS) framework, which utilizes user-centered design, design thinking, and theory-driven behavioral strategies.<sup>32</sup> In our prior study, qualitative insights were gathered from users to inform the features of Medicagent during the Integrate phase.<sup>18</sup> This study

focused on the Design phase to incorporate users' feedback into Medicagent through usability testing. We leveraged the Information-Motivation-Behavioral skills model<sup>33</sup> as we aimed for individuals to be well-informed about hypertension and medications, motivated through reminders and support from their care team, and have the behavioral skills for effective blood pressure management. The model of medication self-management was also used to incorporate the steps involved in successful medication management.<sup>34</sup> We incorporated several distinct features from prior hypertension-related chatbots, including information about refills, medication side effects, and appointment scheduling. Low fidelity prototypes were created through sketches and conversational flow maps. Designs were iteratively discussed and modified with members of the research team. Given the sensitive nature of managing one's medications and health tasks, Medicagent was initially designed for text-based interactions as compared to spoken queries.<sup>35</sup>

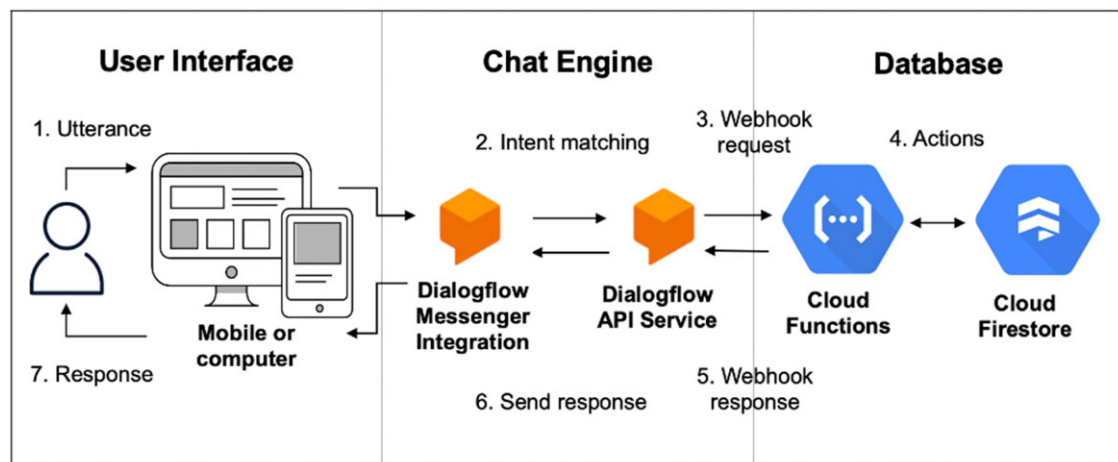
### Chatbot development

The chatbot architecture had 3 primary components: user interface, chat engine, and database.<sup>36</sup> The user interface included a chat dialogue window that could be opened or closed by the user and accessed on a computer or mobile device. Google Cloud Dialogflow natural language understanding platform was used as the chat engine.<sup>37</sup> Dialogue fulfillment tasks were handled using Google Cloud Functions, and Google Cloud Firestore was used as a database. For example, the following steps take the user's utterance or input through the system until a response is returned (Figure 1):

- 1) The user types an utterance or phrase.
- 2) Dialogflow Messenger Integration sends the 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.
- 6) Dialogflow sends the response back to the user.
- 7) The response is returned to the user.

### User interface

The chatbot web interface was compatible with commonly used browsers (Chrome, Edge, Firefox, Safari) and responsive on both computer and mobile devices (iOS and Android). When used on a computer, the chat dialogue window appeared in the lower right side of the screen. On a mobile device, the dialogue window filled the entire screen. We aimed for inclusivity for patients of diverse abilities and used web content accessibility guidelines to guide design decisions, including resizable text, color contrast, and small blocks of text.<sup>38</sup> Medicagent also included visual affordances such as icons, graphics, and buttons, which are typically familiar to those who use text messaging or social media. Figure 2 contains screenshots of the dialogue window.



**Figure 1.** System architecture.\* \*Adapted from: <https://cloud.google.com/dialogflow/es/docs/basics>.

### Chat engine and database

Google Cloud Dialogflow chat engine utilizes 2 algorithms to match natural language utterances to intents: rule-based grammar matching and machine learning (ML) matching.<sup>39</sup> 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). A classification threshold of 0.3 was used for Medicagent. If the highest scoring intent was 0.3 or above, an utterance was matched to the intent. If the score was less than 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 that were created by the study team. Approximately 10-15 training phrases were used for each intent as ML matching expands the list with additional similar phrases. The prototype underwent extensive pilot testing from health professionals (ie, physician, nurse, pharmacist) and informaticians during development. Utterances were iteratively added to the training phrases to improve system performance. Content and images were derived from the U.S. National Library of Medicine’s Pillbox<sup>40</sup> and MedlinePlus.<sup>41</sup>

Appendix S1 contains example training phrases and chatbot responses or prompts for each intent. After an utterance was matched to an intent, parameters were extracted and used to determine the necessary action, such as checking the Firestore database to determine if an appointment time was available. As there was no pre-defined decision logic and the conversation could be driven by the user or system, users were prompted with questions to fill parameters when needed. Regular expressions were used to accept a variety of formats for dates, times, and blood pressure values.

### Study design for usability testing

We used standardized methods for usability studies of mobile health systems to assess effectiveness, efficiency, and satisfaction from the International Organization for Standardization 9241-11.<sup>30,42</sup> Usability testing was conducted virtually during one-on-one Zoom sessions. Effectiveness was measured by the percentage of tasks completed, user task error rate, and system error rate. 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, shortcomings, and experience using Medicagent. Our primary focus was on usability as our prior study assessed acceptance and perceived usefulness of chatbots, where most participants reported positive attitudes toward health chatbots.<sup>18</sup> 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 sample

Eligible participants were adults 18 years and older with hypertension who were prescribed at least one hypertension medication. Participants had to have access to a computer and did not require assistive technology. 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 10 individuals<sup>43</sup> with at least 50% of the following characteristics: 65+ years old, individuals of diverse race/ethnicity, male, education less than college, and taking at least 3 medications.

### Procedures

The usability testing session consisted of 4 components: (1) background questionnaire, (2) representative tasks within Medicagent using hypothetical data, (3) usability questionnaire, and (4) brief semi-structured interview. One researcher (A.C.G.), who had been trained in usability testing and interviewing, conducted the sessions using a testing guide (Appendix S2). The guide and questionnaires were pilot tested with members of the study team.

First, participants completed the background questionnaire on topics of sociodemographics, medical history, experience with technology, health literacy,<sup>44</sup> medication self-efficacy,<sup>45</sup> and barriers to medication adherence.<sup>46</sup> Next, participants were asked to access Medicagent from their computer. We created a study website that contained a collapsed list of tasks on the left side and the chatbot dialogue window on the right side to facilitate the session. Participants were verbally introduced to Medicagent and then asked to complete 5 hypothetical data entry and 5 data retrieval tasks based on

**Panel A:** Medicagent logo. "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!" Input: "Hi".

**Panel B:** "You can also use the chat below to ask questions." "Go ahead and type 'Menu' to see what I can do!" "Here are some options I can help you with." Menu options: Medications, Refills, Health and Blood Pressure Tracking, Appointments, Healthy Resources. Input: "menu".

**Panel C:** "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 🙌"

**Panel D:** "Here are your blood pressures for August 2020." Graph: "Blood Pressure - August 2020". Y-axis: "Blood Pressure (mmHg)" (0-160). X-axis: "Date" (8/1, 8/8, 8/15, 8/22, 8/29). Legend: "Systolic BP" (blue line), "Diastolic BP" (orange line). Buttons: "Add blood pressure", "Share blood pressure".

**Panel E:** "Amlodipine 5mg" "Take 1 each day at 8:00am" [Image of pill] "[View info and side effects of Amlodipine](#)" "Metformin 1,000mg" "Take 2 each day at 8:00am and 9:00pm" "Atorvastatin 40mg" "Take 1 each day at 9:00pm" Buttons: "Add medication", "View medication reminders".

**Panel F:** "View medication refill dates" "Here is your refill information, including the date of the next refill and the number of refills left." List: "Amlodipine 5mg" (Next Refill: 11/15/2020 (2 refills left)), "Metformin 1,000mg" (Next Refill: 11/15/2020 (2 refills left)), "Atorvastatin 40mg" (Next Refill: 12/15/2020 (3 refills left)). Buttons: "Thank you", "You're welcome! What else can I help with?".

**Figure 2.** Chatbot dialogue window. \* \*The user initiates the conversation by typing a variation of "hi" (A). Menu options are triggered by a phrase that contains "menu" (B). 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 (C). Monthly graphs of blood pressure tracking can be viewed and shared with a provider (D). The user can view a list of medications, instructions, side effects (E), and refill information (F).

hypertension and medication self-management processes (Appendix S2). Hypothetical data within Medicagent were the same across all participants, and participants were informed that the tasks did not contain their actual health data. Participants were asked to describe their thoughts, feelings, and actions using concurrent think aloud methods while completing these tasks.<sup>47</sup> The interviewer asked participants to click on a single task at a time, and all tasks were completed in the same order for all participants.

Following task completion, participants completed the SUS questionnaire which is a reliable 10-item questionnaire used to measure the usability of a system.<sup>48</sup> Lastly, participants were asked about their perceptions and experience with Medicagent in a brief semi-structured interview. 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.

## Analysis

Questionnaire scores were calculated following standard scoring methodology from each 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 indicates inadequate health literacy.<sup>44</sup> The Patient-Reported Outcomes Measurement Information System Self-efficacy for Managing Medications and Treatments Short Form 8a (2016) was scored by converting raw scores into T-scores with a mean of 50 (SD = 10) with higher

scores representing greater self-efficacy.<sup>45</sup> The Adherence Starts with Knowledge 12 (ASK-12) scores range from 12 to 60 with higher scores representing greater barriers to adherence.<sup>46</sup> SUS scores range from 0 to 100 with higher scores representing greater usability.<sup>48</sup> Prior approval was obtained for use of the 3-item Brief Health Literacy Screener<sup>44</sup> and ASK-12.<sup>46</sup> Users' clicks, utterances, and task durations were extracted from Google Cloud Dialogflow analytics and verified by watching recorded Zoom sessions. Each session was transcribed verbatim using Zoom transcription and analyzed by a trained qualitative researcher (A.C.G.) using a thematic analysis approach.<sup>49</sup> The reviewer inductively identified and applied thematic content codes to identify the strengths and weaknesses of Medicagent across 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

### Sample characteristics

Ten participants completed usability testing sessions. The average age was 60 years, and half were female, Black, and had at least a college education (Table 1). The majority of participants had been diagnosed with hypertension at least 5 years ago (80%) and took an average of 4 medications. Half had used a chatbot before and reported using Apple Siri, Amazon Alexa, Google Assistant, or customer service chatbots from websites. The majority had scores above the U.S. population average for medication self-efficacy and felt that the greatest barriers to adherence were treatment beliefs (ie, "I feel confident that each one of my medicines will help me").

### Summary of tasks

#### Effectiveness

Nearly all tasks (98%) were successfully completed (Table 2). 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 10 participants, a total of 252 button clicks and 128 utterances were made throughout the testing sessions. Only 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 task. Examples of participants utterances and corresponding intents are shown in Appendix S3.

#### Efficiency

Participants spent an average of 18 min (SD = 10 min) interacting with Medicagent during their usability testing session. Data retrieval tasks were completed faster on average (58 s) as compared to data entry tasks (93 s). Adding a new medication took the most time on average (132 s), whereas scheduling an appointment took the least time on average (44 s). As users progressed through the tasks, data entry task duration decreased, whereas, data retrieval duration somewhat increased (Table 2). This may be due to 2 data retrieval tasks that involved clicking on a hyperlink which opened up a new tab. Several participants found it difficult to navigate between tabs. Overall, data entry tasks had more button clicks and utterances.

**Table 1.** 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	4 (40)
Almost constantly	6 (60)
Device use	
Smartphone	9 (90)
Basic cell phone	1 (10)
Tablet	5 (50)
Computer	10 (100)
Ever used a chatbot	
Yes	5 (50)
No/don't know	5 (50)
Health literacy level <sup>44</sup>	
Adequate	10 (100)
Medication self-efficacy, <sup>45</sup> mean (SE)	51.0 (4.1)
Barriers to adherence, <sup>46</sup> mean (SD)	
Treatment beliefs	9.1 (3.2)
Behaviors	7.2 (2.2)
Inconvenience/forgetfulness	6.7 (2.5)
Total score	23.0 (6.2)

### Satisfaction, strengths, and shortcomings

Medicagent achieved a mean SUS score of 78.8 out of 100 (Table 3). Scores below 50 are generally considered not acceptable, 50 to 70 as marginal, and above 70 acceptable.<sup>50,51</sup> Participants reported a number of strengths and shortcomings of Medicagent (Table 4). Overall, most reported Medicagent was easy and enjoyable to use. Several felt it became simpler to use as the tasks proceeded. Nearly all had positive attitudes towards the visuals, including the images of medications and blood pressure charts.

**Table 2.** Task summary (mean, SD unless noted otherwise).

Task	Duration in seconds	Button clicks	Utterances	User errors (error rate) <sup>a</sup>	Chatbot errors (error rate) <sup>b</sup>
1. Find the list of current medications. <sup>c</sup>	46.1 (25.4)	1 (0)	0 (0)	0	0
2. Find 2 of the side effects of Amlodipine. <sup>c</sup>	69.1 (25.3)	3 (1)	1 (1)	0	0
3. Add 10 mg of Lisinopril every day at noon to your medication list. <sup>d</sup>	131.6 (84.1)	2 (1)	3 (1)	1 (10%) <sup>e</sup>	4 (12.9%) <sup>f</sup>
4. Update your medication reminder schedule for Amlodipine to remind you to take it on weekends at 10 AM. <sup>d</sup>	120.7 (52.6)	4 (1)	2 (1)	1 (10%) <sup>g</sup>	1 (4.5%) <sup>f</sup>
5. View how many refills are left and the date of your next refill for Amlodipine. <sup>c</sup>	60.8 (36.1)	1 (1)	1 (1)	0	0
6. Find your blood pressure values for the month of August 2020. <sup>c</sup>	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. <sup>d</sup>	100.1 (66.8)	2 (1)	3 (1)	0	6 (23.1%) <sup>f,h</sup>
8. Share your blood pressure measurements for the month of August 2020 with Dr. Smith. <sup>d</sup>	69.3 (29.6)	4 (1)	1 (1)	0	0
9. Schedule an appointment with Dr. Smith for Tuesday, November 17th at 1 PM. <sup>d</sup>	43.8 (24.1)	4 (2)	1 (1)	0	0
10. Find a healthy dinner recipe. <sup>c</sup>	60.2 (28.8)	2 (1)	1 (1)	0	0
Total (mean, SD)	75.6 (51.5)	2.5 (1.7)	1.3 (1.2)	2 (2.0%) <sup>a</sup>	11 (8.6%) <sup>b</sup>

<sup>a</sup> User error rate was calculated per task by: number of participants who did not successfully complete the task/total number of participants.

<sup>b</sup> Chatbot error rate was calculated per task by: number of unrecognized utterances/total number of utterances.

<sup>c</sup> Data retrieval task.

<sup>d</sup> Data entry task.

<sup>e</sup> Type of user error: did not confirm medication was added.

<sup>f</sup> Type of unrecognized error: date/time format.

<sup>g</sup> Type of user error: inputted incorrect medication.

<sup>h</sup> Type of unrecognized error: blood pressure format.

**Table 3.** System usability scale scores.<sup>a</sup>

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)
<b>Total SUS score</b>	<b>78.8 (15.9)</b>

<sup>a</sup> 1 = strongly disagree and 5 = strongly agree.

In terms of 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, some participants described a hamburger-like icon that included drop-down options of common tasks. Additionally, many felt recovering from errors was challenging and suggested having a back button in the chat or more instructive error messages. Several thought the chatbot could be improved by adding a persona of a health professional, such as a medical avatar that could interact with them. A few also wanted to interact with Medicagent through voice or a combination of voice and text.

## Discussion

### Principal findings

We designed, developed, and evaluated the usability of a hypertension self-management chatbot prototype,

Medicagent. Participants spent an average of 18 min (SD = 10 min) completing 10 self-management tasks, which is similar to other health chatbot usability interactions ranging between 10 and 25 min.<sup>10</sup> A total of 98% of the tasks were successfully completed, and most felt there was a short learning curve. Medicagent received an average SUS score of 78.8 which demonstrates acceptable usability. In comparison to other text-based health chatbots, mean SUS scores ranged from 79.9 to 88.2.<sup>52–57</sup> Most of these studies evaluated usability of mental health chatbots, and the participant characteristics varied from our study (ie, 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 but not always found in chatbot interfaces. Many desired a health professional identity to be embodied in the

**Table 4.** Themes for usability strengths and shortcomings of Medicagent.

Strengths	Representative quotes
Easy to use	<i>"There was a short learning curve, and it was easier as the tasks went on as I remembered what the menu items were."</i>
Enjoyable	<i>"It even gives me a little motivation to continue to exercise and take my Lisinopril. . .this is quite fun."</i>
Visuals	<i>"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."</i>
<b>Shortcomings</b>	
Menu button	<i>"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."</i>
Error recovery	<i>"It seems like a back button is needed. I made an error and had to go all the way through the questions again." "Once I made an error, since it's the first time I've used this, I didn't know how to get back. . .I've found when I'm doing computer things, if it's not so clear and I get lost, I normally just quit."</i>
Navigating between tabs	<i>"Going to so many different screens sometimes can confuse me and then getting back to the chatbot screen was not that easy."</i>
Persona	<i>"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."</i>

chatbot. We identified 3 main components from qualitative feedback that may facilitate usability of chatbots for self-management: interaction flexibility, graceful degradation (ie, the ability of the system to tolerate failures),<sup>58</sup> and a medical professional persona.

### Interaction flexibility

We observed 2 primary types of interaction styles driven by buttons or utterances. Participants who mainly used buttons began tasks by clicking on a menu button, while those who used utterances started tasks using the text input box. A few completed the first few tasks using the button approach and then switched to utterances as tasks progressed. Prior studies have found that users prefer to have the option to input free-text or use buttons.<sup>59</sup> However, unrecognized free-text responses of health information may pose potential patient safety risks if appropriate measures and human handoffs are not in place.<sup>9,60</sup> Confirming user input may help mitigate some of these concerns, and several participants found the confirmations to validate their inputs helpful. For those who preferred navigating from the menu, a menu icon at the top was desired to select from a dropdown list. Several participants also felt that a back button would be useful to navigate to a prior point in the conversation instead of re-querying. Some wanted the flexibility to communicate through both voice and text, which may be due to half of participants in this study reporting previous use of voice assistants. Careful considerations should be made for the development of interactive systems with spoken queries when handling sensitive health information. Overall, our findings suggest that users have diverse interaction preferences for self-management chatbots, which may be addressed through various visual cues and input mechanisms.

### Graceful degradation

Participants inputted a variety of phrases, and some described how the default fallback messages were too generic. Handling unrecognized utterances gracefully is particularly important for chatbots because people generally perceive robots as intelligent and may have less tolerance for mistakes.<sup>61,62</sup> Erroneous agents are also perceived as less reliable which negatively

affects task performance.<sup>61,62</sup> Currently, most chatbots are still limited by their inability to hold meaningful conversations and personalized interactions.<sup>14</sup> 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 frustration.<sup>63</sup> 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.<sup>21</sup>

### Medical professional persona

Several participants desired 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 of the system. 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.<sup>64</sup> Prior studies of ECAs have simulated a health provider for health coaching and reviewing hospital discharge materials.<sup>11,20,65,66</sup> Overall, these have been received positively by patients and increased adherence to treatment regimens.<sup>11,20,65,66</sup> It is possible that some participants were used to interacting with chatbots by their names, such as "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 rapport and engagement with the chatbot.

### Implications for healthcare and research

This research has several implications for the field of chatbots for health. We provide an overview of our design process and highlight several components that may enrich the user experience of health-related chatbots. This may provide some insights for chatbot development in the context of other chronic conditions requiring self-management. We used a labor-intensive process of iteratively training the chat engine through pilot testing, assessing the chat logs for unrecognized utterances, and then adding utterances and intents to the

training data. Similar manual processes have been reported in chatbot development studies to fine-tune the system.<sup>67</sup> Zand et al<sup>68</sup> used natural language processing to develop a chatbot knowledge base by categorizing electronic messages between patients and providers. Comparable approaches may be useful to inform self-management knowledge bases, which could be used during chatbot development and tailored for cultural and linguistic differences. Most patient-provider communication and self-management knowledge bases are not publicly available due to confidentiality issues.<sup>69–71</sup> 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<sup>72,73</sup> that may also be a useful starting point in the design and development of future chatbots. As with all technologies, careful consideration should be made to promote inclusive designs and ensure representative patient groups are involved in the design and algorithm training.

For our prototype, we used hypothetical data and did not include protected health information (PHI), which would require a Google Cloud Business Associate Agreement to ensure Health Insurance Portability and Accountability Act (HIPAA) compliance.<sup>74</sup> Several voice apps are currently operating under Amazon Alexa's HIPAA eligible environment to allow users to query their blood sugar measurements or check the status of home delivery prescriptions.<sup>75</sup> As conversational interfaces are becoming more widespread and connecting to various APIs, confidentiality and privacy of PHI should be carefully monitored. The increasing focus on access to patient-level data and APIs through the 21st Century Cures Act<sup>76</sup> and Interoperability and Patient Access<sup>77</sup> final rules may spur additional innovations in personalized self-management and communication tools.

### Limitations

The sample was limited to individuals with a computer from a single geographic region in the Southeast, and participants had adequate health literacy, high levels of medication self-efficacy, and high educational attainment. Prior research suggests that our sample size is adequate for usability testing,<sup>43</sup> but we may not be able to draw definitive conclusions about the strengths and shortcomings of self-management chatbots across diverse populations. These factors may limit the generalizability of our findings. In addition, concurrent think aloud methods are valuable for real-time feedback, although 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. Due to the COVID-19 pandemic, usability testing was conducted remotely over Zoom on participants' personal computers. Although the website that contained Medicagent was consistent across participants, there were variances in Internet speed and potential differences in hardware. However, conducting remote usability testing allowed us to observe interactions within a user's natural environment at home where self-management tasks would typically take place.

### Conclusion

In the emergent field of health 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 were highlighted as design features that could facilitate usability and navigation within chatbots for self-management. Assessing usability is an important step towards understanding how patients interact with chatbots to achieve self-management tasks, and this research may help inform how designers can improve patient experience and engagement of conversational systems. 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.

### Acknowledgments

We are grateful to Dr. Jaime Arguello and Dr. Yue Wang for their contributions to this work. We would also like to thank the participants for taking part in this study.

### Author contributions

A.C.G. and A.E.C. were responsible for overall study concept and design. All authors contributed to the study design, questionnaire development, testing session guide, manuscript writing, and revision. A.C.G. was primarily responsible for data collection, analyses, and initial draft of the manuscript. All authors reviewed, edited, and approved the final version.

### Supplementary material

[Supplementary material](#) is available at *JAMIA Open* online.

### Funding

Research reported in this publication was supported by the National Institutes of Health's National Center for Advancing Translational Sciences through Grant Number UL1TR002489. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. A.C.G. acknowledges funding support by National Institutes of Health National Library of Medicine training grant 5T15LM012500-04, while at the University of North Carolina at Chapel Hill when this work was completed. A.C.G. is currently supported by a VA Advanced Fellowship in Medical Informatics. The opinions expressed are those of the authors and not necessarily those of the Department of Veterans Affairs or those of the United States Government.

### Conflict of interest

A.E.C. is now employed by Google, LLC and has equity in the company. However, this work was completed prior to employment at Google when A.E.C. was at the University of North Carolina at Chapel Hill School of Medicine. No resources or funding were provided by Google for this research. S.C.B. reports grants from the NIH, Merck, Pfizer, Gordon and Betty Moore Foundation, Retirement Research Foundation for Aging, Lundbeck, and Eli Lilly via her institution and personal fees from Sanofi, Pfizer, University of Westminster, Lundbeck, and Luto outside the submitted work.



## Data availability

The data underlying this article will be shared upon reasonable request to the corresponding author.

## References

- Benjamin EJ, Muntner P, Alonso A, *et al.*; On behalf of the American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2019 update: a report from the American Heart Association. *Circulation*. 2019;139(10):e56-e528.
- Centers for Disease Control and Prevention. *Hypertension Cascade: Hypertension Prevalence, Treatment and Control Estimates Among US Adults Aged 18 Years and Older Applying the Criteria From the American College of Cardiology and American Heart Association's 2017 Hypertension Guideline—NHANES 2015-2018*. Atlanta, GA: US Department of Health and Human Services; 2020.
- 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.
- 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. *Curr Atheroscler Rep*. 2017;19(3):12.
- Xiong S, Berkhouse H, Schooler M, *et al.* Effectiveness of mHealth interventions in improving medication adherence among people with hypertension: a systematic review. *Curr Hypertens Rep*. 2018;20(10):86.
- 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):801-817.
- 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.
- 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.
- 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.
- Griffin AC, Xing Z, Khairat S, *et al.* Conversational agents for chronic disease self-management: a systematic review. *AMIA Annu Symp Proc*. 2020:504-513.
- Bickmore T, Giorgino T. Health dialog systems for patients and consumers. *J Biomed Inform*. 2006;39(5):556-571.
- Uohara MY, Weinstein JN, Rhew DC. The essential role of technology in the public health battle against COVID-19. *Popul Health Manag*. 2020;23(5):361-367.
- Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine. *Commun ACM*. 1966;9(1):36-45.
- McTear M. *Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots*. San Rafael, CA: Morgan & Claypool Publishers; 2020.
- 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. *Ann Behav Med*. 2012;43(1):62-73.
- Persell SD, Peprah YA, Lipszko 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.
- Echezarra L, Pereira J, Saracho R. TensioBot: a chatbot assistant for self-managed in-house blood pressure checking. *J Med Syst*. 2021;45(4):54.
- Griffin AC, Xing Z, Mikles SP, *et al.* Information needs and perceptions of chatbots for hypertension medication self-management: a mixed methods study. *JAMIA Open*. 2021;4(2):ooab021.
- Giorgino T, Azzini I, Rognoni C, *et al.* Automated spoken dialogue system for hypertensive patient home management. *Int J Med Inform*. 2005;74(2-4):159-167.
- 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.
- Bickmore TW, Caruso L, Clough-Gorr K. Acceptance and usability of a relational agent interface by urban older adults. In: *CHI '05 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery; April 2–7, 2005:1212-1215; Portland, OR.
- Tsiourti C, Joly E, Wings C, *et al.* Virtual assistive companions for older adults: qualitative field study and design implications. In: *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering; 2014;57-64; Oldenburg, Germany.
- 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.
- 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.
- 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.
- 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.
- Baptista S, Wadley G, Bird D, *et al.*; My Diabetes Coach Research Group. 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.
- 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.
- Yen P, Bakken S. Review of health information technology usability study methodologies. *J Am Med Inform Assoc*. 2012;19(3):413-422.
- International Organization for Standardization. Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs)—Part 11: Guidance on Usability. 9241-11.1998. Accessed January 20, 2021. <https://www.iso.org/standard/16883.html>.
- Carayon P, Hoonakker P. Human factors and usability for health information technology: old and new challenges. *Yearb Med Inform*. 2019;28(1):71-77.
- 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.
- Fisher WA, Fisher JD, Harman J. The information-motivation-behavioral skills model: a general social psychological approach to understanding and promoting health behavior. In: Suls J, Wallston KA, eds. *Blackwell Series in Health Psychology and Behavioral Medicine. Social Psychological Foundations of Health and Illness*. Blackwell Publishing; 2003;82-106.
- Bailey SC, Oramasionwu CU, Wolf MS. Rethinking adherence: a health literacy—informed model of medication self-management. *J Health Commun*. 2013;18(Suppl 1):20-30.
- Guy I. Searching by talking: analysis of voice queries on mobile web search. In: *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Association for Computing Machinery; 2016; Pisa, Italy.

36. Google Cloud Dialogflow. API Interactions. 2021. Accessed December 9, 2021. <https://cloud.google.com/dialogflow/es/docs/basics>.
37. Google Cloud. Dialogflow. Accessed December 19, 2018. <https://dialogflow.com/>.
38. World Wide Web Consortium. Web Content Accessibility Guidelines Overview. 2020. Accessed November 11, 2020. <https://www.w3.org/WAI/standards-guidelines/wcag/>.
39. Google Cloud Dialogflow. Intent Matching. 2020. Accessed December 1, 2020. <https://cloud.google.com/dialogflow/es/docs/intents-matching#algo>.
40. U.S. National Library of Medicine. Pillbox. 2019. Accessed November 1, 2019. <https://pillbox.nlm.nih.gov/>.
41. U.S. National Library of Medicine. MedlinePlus. 2019. Accessed November 1, 2019. <https://medlineplus.gov/>.
42. Georgsson M, Stagers 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.
43. 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.
44. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. *Family Medicine*. 2004;36(8):588-594.
45. Patient-Reported Outcomes Measurement Information System. Self-Efficacy for Managing Medications and Treatments—Short Form 8a. 2016. Accessed August 5, 2019. [http://www.healthmeasures.net/administrator/components/com\\_instruments/uploads/PROMIS%20SF%20v1.0%20-%20Self-Effic-ManagMeds%208a\\_8-5-2016.pdf](http://www.healthmeasures.net/administrator/components/com_instruments/uploads/PROMIS%20SF%20v1.0%20-%20Self-Effic-ManagMeds%208a_8-5-2016.pdf).
46. Matza LS, Park J, Coyne KS, et al. Derivation and validation of the ASK-12 adherence barrier survey. *Ann Pharmacother*. 2009;43(10):1621-1630.
47. Lewis C. *Using the "Thinking Aloud" Method in Cognitive Interface Design*. NY: IBM Watson Research Center; 1982.
48. Brooke J. *SUS—A Quick and Dirty Usability Scale*. *Usability Evaluation in the Industry*. London, England. Taylor & Francis; 1996:189-194.
49. Guest G, MacQueen KM, Namey EE. *Applied Thematic Analysis*. Thousand Oaks, CA: Sage Publications, Inc.; 2012.
50. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud*. 2009;4(3):114-123.
51. Brooke J. SUS: a retrospective. *J Usability Stud*. 2013;8(2):29-40.
52. Schroeder J, Wilkes C, Rowan K, et al. Pocket skills: a conversational mobile web app to support dialectical behavioral therapy. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery; 2018:1-15; Montreal, QC, Canada.
53. Cameron G, Cameron D, Megaw G, et al. Assessing the usability of a chatbot for mental health care. In: *5th International Conference on Internet Science*. Springer Verlag; 2018:121-132; St. Petersburg, Russia.
54. Islam MN, Khan SR, Islam NN, et al. A mobile application for mental health care during COVID-19 pandemic: development and usability evaluation with system usability scale. In: Suhaili WSH, Siau NZ, Omar S, Phon-Amuisuk S, eds. *Computational Intelligence in Information Systems*. Springer, Cham; 2021:1321.
55. Holmes S, Moorhead A, Bond R, et al. Usability testing of a healthcare chatbot: can we use conventional methods to assess conversational user interfaces? In: *Proceedings of the 31<sup>st</sup> European Conference on Cognitive Ergonomics*. Association for Computing Machinery; 2019:207-214; Belfast, United Kingdom.
56. da Silva Lima Roque G, Roque de Souza R, Araújo do Nascimento JW, et al. Content validation and usability of a chatbot of guidelines for wound dressing. *Int J Med Inform*. 2021;151:104473.
57. Issom DZ, Hardy-Dessources MD, Romana M, et al. Toward a conversational agent to support the self-management of adults and young adults with sickle cell disease: usability and usefulness study. *Front Digit Health*. 2021;3:600333.
58. Shelton CP, Koopman P, Nace W. A framework for scalable analysis and design of system-wide graceful degradation in distributed embedded systems. In: *Proceedings of the Eighth International Workshop on Object-Oriented Real-Time Dependable Systems*. IEEE; 2003:156-163; Guadalajara, Mexico.
59. Budiu R. The User Experience of Chatbots. 2018. Accessed January 1, 2021. <https://www.nngroup.com/articles/chatbots/>.
60. 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.
61. 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)*. IEEE; 2016:501-506; New York, NY.
62. Toader D, Boca G, Toader R, et al. The effect of social presence and chatbot errors on trust. *Sustainability*. 2019;12(1):256.
63. Ceaparu I, Lazar J, Bessiere K, et al. Determining causes and severity of end-user frustration. *Int J Hum Comput Interact*. 2004;17(3):333-356.
64. Cassell J, Sullivan J, Preost S, et al. *Embodied Conversational Agents*. Cambridge, UK: MIT Press; 2000.
65. Bickmore TW, Caruso L, Clough-Gorr K, et al. It's just like you talk to a friend' relational agents for older adults. *Interact Comput*. 2005;17(6):711-735.
66. 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.
67. Preininger AM, South B, Heiland J, et al. Artificial intelligence-based conversational agent to support medication prescribing. *JAMIA Open*. 2020;3(2):225-232.
68. 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.
69. Staples S. *The Discourse of Nurse-Patient Interactions: Contrasting the Communicative Styles of U.S. and International Nurses*. Philadelphia, PA: John Benjamins Publishing; 2015.
70. Thomas J, Short M. *Using Corpora for Language Research*. London, UK: Longman Pub Group; 1996.
71. Adolphs S, Brown B, Carter R, et al. Applying corpus linguistics in a health care context. *J Appl Linguist*. 2004;1(1):9-28.
72. Microsoft Azure. Health Bot Overview. 2020. Accessed July 1, 2020. <https://docs.microsoft.com/en-us/azure/health-bot/>.
73. Mehta N, Petersen K, To W, et al. AWS for Industries: Building Clinically-Validated Conversational Agents to Address Novel Coronavirus. 2020. Accessed January 1, 2021. <https://aws.amazon.com/blogs/industries/building-clinically-validated-conversational-agents-to-address-novel-coronavirus/>.
74. Google Cloud. Google Cloud Services that are in Scope for HIPAA. 2020. Accessed December 20, 2020. <https://cloud.google.com/security/compliance/hipaa-compliances>.
75. Amazon Web Services. Introducing New Alexa Healthcare Skills. Accessed April 14, 2019. <https://developer.amazon.com/blogs/alexa/post/ff33dbc7-6cf5-4db8-b203-99144a251a21/introducing-new-alexa-healthcare-skills>.
76. Department of Health and Human Services. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program. 2020. Accessed June 1, 2021. <https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification>.
77. Centers for Medicare & Medicaid Services. Interoperability and Patient Access Final Rule. Accessed June 1, 2021. <https://www.cms.gov/Regulations-and-Guidance/Guidance/Interoperability/index#CMS-Interoperability-and-Patient-Access-Final-Rule>.