International Journal of Medical Informatics 192 (2024) 105640

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

A conversational agent for enhanced Self-Management after cardiothoracic surgery

Ana Martins [a,b,*], Luís Velez Lapão [a,c,d], Isabel L. Nunes [a,c], Ana Paula Giordano [b,e], Helena Semedo [g], Clara Vital [g], Raquel Silva [h], Pedro Coelho [d,g], Ana Londral [b,d,f]

[a] UNIDEMI, Department of Mechanical and Industrial Engineering, Nova School of Science and Technology, Caparica 2829-516, Portugal
[b] Value for Health CoLAB, Lisboa 1150-190, Portugal
[c] Laboratório Associado de Sistemas Inteligentes, Escola de Engenharia Universidade do Minho, Campus Azurém, 4800-058 Guimarães, Portugal
[d] Comprehensive Health Research Center, Nova Medical School, Lisboa 1169-056, Portugal
[e] CUBE - CATÓLICA-LISBON Research Unit in Business and Economics, Portugal
[f] Department of Physics, Nova School of Science and Technology, Caparica 2829-516, Portugal
[g] Hospital de Santa Marta, Centro Hospitalar Universitário Lisboa Central, Rua de Santa Marta, 50, 1169-023 Lisboa Portugal
[h] NOVA CLUNL – Nova University Lisbon, Colégio Almada Negreiros, 1099-032 Lisboa, Portugal

ARTICLE INFO

Keywords:
Co-design
Conversational Agents
Personalization
Self-management
Cardiothoracic Surgery
Health

ABSTRACT

Background: Enhanced self-management is crucial for long-term survival following cardiothoracic surgery.
Objectives: This study aimed to develop a conversational agent to enhance patient self-management after cardiothoracic surgery.
Methodology: The solution was designed and implemented following the Design Science Research Methodology. A pilot study was conducted at the hospital to assess the feasibility, usability, and perceived effectiveness of the solution. Feedback was gathered to inform further interactions. Additionally, a focus group with clinicians was conducted to evaluate the acceptability of the solution, integrating insights from the pilot study.
Results: The conversational agent, implemented using a rule-based model, was successfully tested with patients in the cardiothoracic surgery unit (n = 4). Patients received one month of text messages reinforcing clinical team recommendations on a healthy diet and regular physical activity. The system received a high usability score, and two patients suggested adding a feature to answer user prompts for future improvements. The focus group feedback indicated that while the solution met the initial requirements, further testing with a larger patient cohort is necessary to establish personalized profiles. Moreover, clinicians recommended that future iterations prioritize enhanced personalization and interoperability with other hospital platforms. Additionally, while the use of artificial generative intelligence was seen as relevant for content personalization, clinicians expressed concerns regarding content safety, highlighting the necessity for rigorous testing.
Conclusions: This study marks a significant step towards enhancing post-cardiothoracic surgery care through conversational agents. The integration of a diversity of stakeholder knowledge enriches the solution, grants ownership and ensures its sustainability. Future research should focus on automating message generation and delivery based on patient data and environmental factors. While the integration of artificial generative intelligence holds promise for enhancing patient interaction, ensuring the safety of its content is essential.

1. Introduction

Cardiovascular disease (CVD) significantly impacts many Europeans [1], and managing modifiable risk factors is crucial for secondary prevention [2]. Cardiothoracic surgery helps reduce mortality and improve quality of life for CVD patients [3]. Post-surgery, effective self-management is essential for long-term survival and reduced need for further interventions [3,4]. Self-management involves daily health-promoting activities and active patient participation in treatment.

In recent years, digital health interventions (DHI) have emerged as

* Corresponding author at: UNIDEMI, Department of Mechanical and Industrial Engineering, Nova School of Science and Technology, Caparica 2829-516, Portugal.

E-mail address: acd.martins@campus.fct.unl.pt (A. Martins).

https://doi.org/10.1016/j.ijmedinf.2024.105640
Received 23 July 2024; Received in revised form 18 September 2024; Accepted 22 September 2024
Available online 24 September 2024

promising tools to enhance self-management by offering scalable and accessible solutions [5]. However, for DHIs to achieve widespread adoption and effectiveness, they must be designed following human-centered approach and undergo iterative optimization [5]. Key methods for improving DHIs include interviews, co-design, prototyping, and user testing, with feasibility studies ensuring acceptability and integration [5].

Despite this, there is limited evidence on the use of conversational agents for post-surgery self-management. To address this gap, our project developed an artifact using existing literature and a multidisciplinary team to enhance self-management for cardiothoracic surgery patients. The project has three main contributions:

1. Understanding factors influencing post-surgery self-management from both clinical and patient perspectives;
2. Integrating stakeholder feedback using qualitative and quantitative methods to define and design a solution;
3. Pilot testing a conversational agent to evaluate its usability, satisfaction, and perceived effectiveness, incorporating feedback from patients and clinical teams.

This study is part of the CardioFollow. AI project at the Cardiothoracic Surgery Department of Hospital de Santa Marta in Lisbon. It includes a remote patient monitoring system (RPM) for acute recovery [6] and proposes a conversational agent to support recovery beyond the initial monitoring period.

2. Methodology

The Design Science Research (DSR) methodology was used to develop an innovative and successful artifact [7]. The process includes six main iterative activities [8], as shown in Fig. 1. The following sections provide detailed explanations of each methodology used.

2.1. Step 1: Problem and motivation

This step was exploratory and aimed to describe the context in which the solution is to be used, the needs of the main stakeholders, and the target population (patients after cardiothoracic surgery), to find a problem that has not yet been solved in the literature and to motivate the research efforts.

2.1.1. Population Characterization: Retrospective data analysis

Retrospective anonymized clinical data collected since 2010 at the hospital. The dataset includes 105 variables, describing patients' demographics and pre-surgery conditions, conducted procedures, length of stay, and in-hospital post-surgery complications. It follows the European Association for Cardio-Thoracic Surgery (EACTS) Adult Cardiac Database fields. Due to study protocol, the dataset cannot be made publicly available.

The Shapiro-Wilk Test was used to check for normal distribution of numerical variables [9]. Normally distributed variables were described using the mean and standard deviation; otherwise, they were represented by the median and interquartile range. Categorical variables were described using frequency and percentage. Data analysis was performed using Python [10].

2.1.2. Digital health interventions: Literature review

A literature review was performed to understand the body of knowledge regarding the DHIs for patients with chronic diseases. A literature search was performed to find studies from the 1st of January 2017 to the 26th of March 2022, in PubMed, Scopus, Web of Science, and CENTRAL. Appendix A further explains the methodology.

2.1.3. In-depth semi-structured interview with the clinical team

In-depth semi-structured interviews [11] were selected to provide an overview of factors influencing patient recovery [2,12–15]. Structured elements were included to simplify collecting clinicians' perspectives. Recruitment occurred after a June 2022 project meeting, followed by an email detailing the interview's content to eight participants. The interview guide, tested with a multidisciplinary team, included questions on modifiable risk factors, psychosocial factors, hospital context, and recovery suggestions (Appendix B). LimeSurvey was used to conduct the interviews [16].

2.1.4. Semi-structured interview with patients

Interviews were conducted via phone in July 2022 with 9 patients. Two nurses helped develop the interview guide, which a health literacy specialist reviewed. Interviews began after obtaining patient consent and aimed to understand their post-cardiothoracic surgery experience, well-being factors, barriers and facilitators to following clinical recommendations, and willingness to use digital tools. Interviews were transcribed verbatim, coded using Taguette with both deductive and inductive approaches [17], as reported elsewhere [18]. The full interview guide is in Appendix C.

2.2. Step 2: Defining the objectives for a solution

This step involved compiling previous information and using a prototyping session with three clinical team members to collaboratively set objectives for the solution, focusing on target behavior, timing, dosing, and communication channel. Technical specifications were designed to ensure the solution can scale and generalize for use in other contexts.

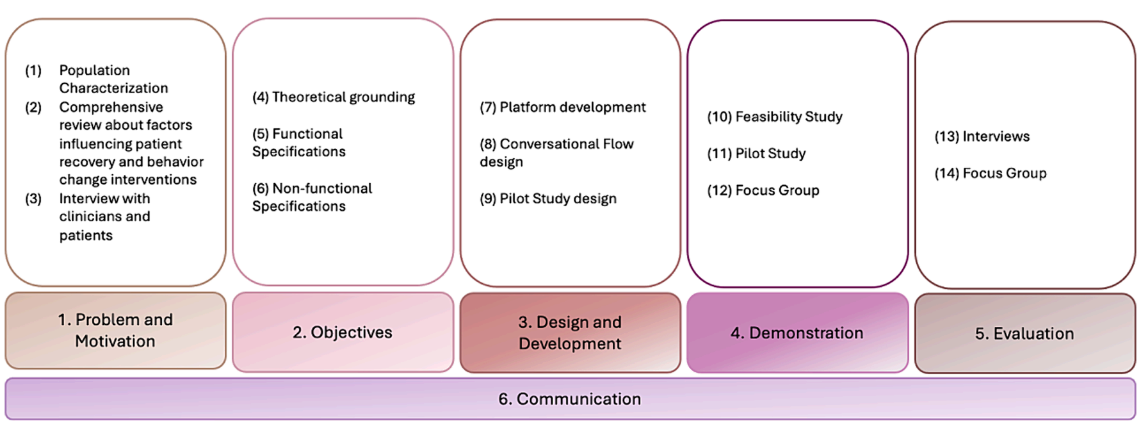

Fig. 1. First DSR iteration: all the activities developed during this 6-step cycle.

2.3. Step 3: Design and development

In this step, the platform, intervention, and content were designed according to objectives from step 2. The Django framework was used to implement the platform for message conveyance and visualization [19]. Twilio was used for messaging, and Docker for creating a deployable image on an AWS EC2 instance. Software versions were saved on GitHub. Functionality testing was conducted by an independent person. Fig. 2 shows the platform's architecture.

A conversational flow design session [20] with two nurses involved in the self-management process occurred to select the conversation flow. The text messages were tested in a prior study [21]. Later, they were translated to Portuguese and adapted to our context.

2.4. Step 4 & 5: Demonstration and evaluation

A feasibility study with 7 researchers assessed the artifact's feasibility and usability. Subsequently, the artifact was piloted at a hospital with 4 patients after the 3-month RPM period. These patients were recruited by their nurse via phone call. Those who agreed to participate had an in-person meeting with their nurse to explain the system and obtain formal consent. A questionnaire (Appendix D) collected patient preferences. Acceptability was assessed by the ratio of patients who agreed to participate to those invited. Retention was calculated as the ratio of participants who completed the one-month duration without suspending message delivery to those who initially agreed. Engagement was measured by the ratio of answers provided to expected answers. Additionally, patient experience was collected via phone calls, and system usability was assessed using the SUS, along with perceived efficacy in improving diet and physical activity, and potential improvements.

A focus group with nursing team members responsible for post-surgery follow-up was conducted by two moderators (AM and AG). Results of the pilot study were presented, and satisfaction, improvement suggestions, and acceptance were assessed using the UTAUT [22]. Sessions were audio-recorded with consent, transcribed, and analyzed using deductive and inductive approaches. Appendix E contains the focus group guide.

2.5. Step 6: Communication

Involved all the conferences [18,23] and research articles [24] published within this project.

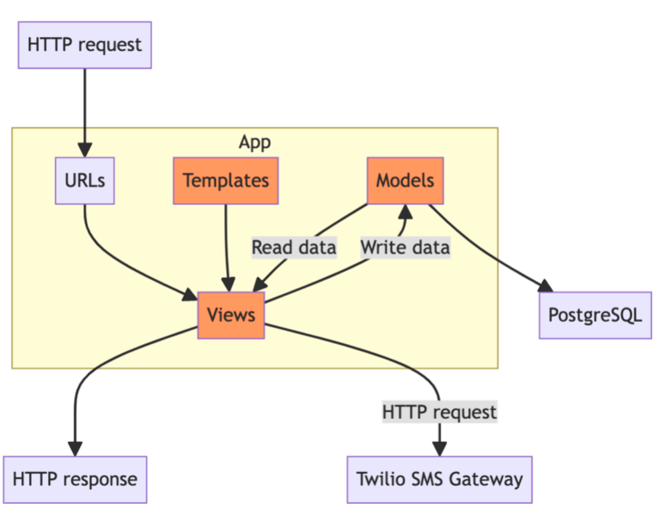

Fig. 2. Model view template architecture for each app.

3. Results

3.1. Problem definition and motivation

3.1.1. Demographics

The study's demographic shows that most participants are retirees over 60 years old, with low education levels and modest incomes, as detailed in Table 1.

3.1.2. Patient pathway

Following interviews with the clinical team, the patient pathway was fully understood (Fig. 3). Over the course of a year, the patient undergoes appointments with the cardiothoracic surgeon at the hospital, supplemented by follow-up assessments that occur before surgery and at 3 days, 1 month, 3 months, 6 months, and 1-year intervals after surgery. Additionally, during the first three months after surgery, the patient may be telemonitored.

3.1.3. Modifiable risk factors

Interviews with the clinical team highlighted a strong agreement on the critical role of modifiable risk factors in patient recovery, as shown in Fig. 4. Nurses emphasize these factors in follow-up and intervene when patients don't follow recommendations. Key areas of focus include managing diabetes, adhering to medication, and promoting physical activity, all of which are crucial for effective disease management (Fig. 4). In Fig. 5, can be observed that the clinical team does not think that patients consistently improve their modifiable risk factors. Patients are perceived to often struggle with managing dyslipidemia, exercising regularly, controlling weight, moderating alcohol intake, and maintaining a healthy diet, as depicted in Fig. 5.

3.1.4. Factors influencing self-management: Barriers and facilitators

In the point of view of the clinical team, internal factors such as anxiety, depression, lack of motivation, low health literacy, fear, and pain are perceived as factors that impair patient's ability to self-manage. External factors arise for disadvantaged patients with low socioeconomic status, limited social support, lack of specialized medical care in

Table 1
Population demographics.

	N	*Med(x) (IQR)*	*Mode (x)*	*%*
Age	*1762*	*67,79 (58,56–76,11)*	*61,95*	–
Sex				
Male	*1126*	–	–	*63,90*
Female	*636*	–	–	*36,10*
BMI	*2515*	*26,72 (24,0–29,43)*	*27,55*	–
Level of Education [25]		–	–	
Primary education 1st cycle	*936*	–	–	*53,12*
Primary education 2ndcycle	*212*	–	–	*12,03*
Primary education 3rd cycle	*194*	–	–	*11,01*
Secondary education	*149*	–	–	*8,46*
Higher education	*180*	–	–	*10,22*
None	*91*	–	–	*5,16*
Work situation				
Retired	*1166*	–	–	*66,17*
Employed	*384*	–	–	*21,79*
Unemployed	*85*	–	–	*4,82*
Unable to work (medical leave)	*60*	–	–	*3,41*
Student	*4*	–	–	*0,23*
Other (domestic work or agriculture)	*63*	–	–	*3,58*
Income				
< 350 €	*200*	–	–	*11,35*
350 – 900 €	*681*	–	–	*38,65*
901 – 2000 €	*375*	–	–	*21,28*
2001 – 3500 €	*64*	–	–	*3,63*
> 3500 €	*14*	–	–	*0,79*
Unknown	*428*	–	–	*24,29*

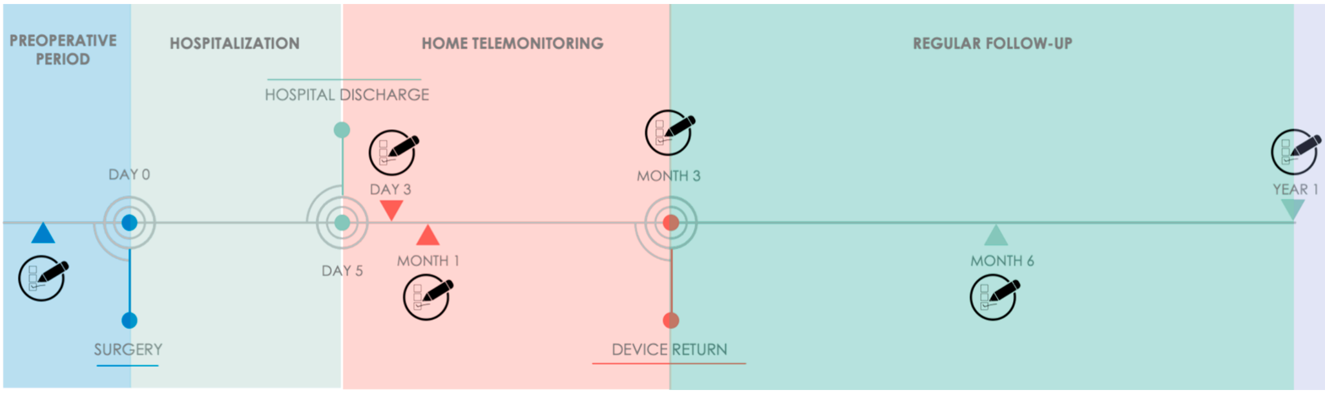

Fig. 3. Current patient pathway at Santa Marta Hospital.

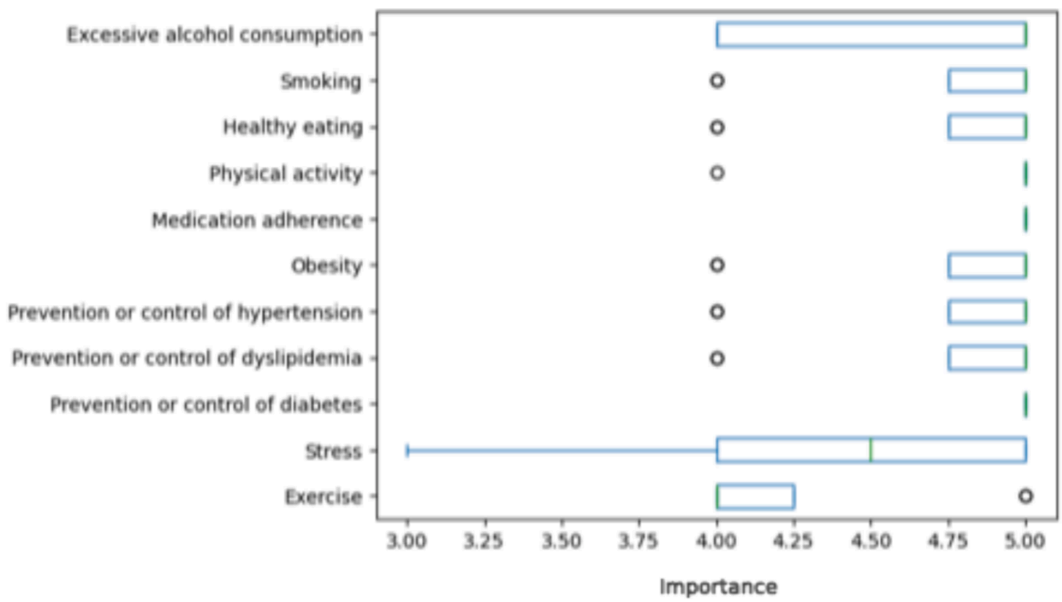

Fig. 4. Median and interquartile perceived importance of modifying risk factors by the clinical team.

Fig. 5. Perceived frequency by which patients do activities to improve the modifying risk factors by the clinical team.

their area, and restricted mobility, impacting their recovery journey.

The patients' perspectives on the barriers and facilitators for self-management have been published elsewhere [18].

3.1.5. Problem definition

The clinical team identified rehospitalization one year after surgery as a key issue. It was hypothesized that patients' self-management skills decline as they feel better, leading to potential problems. To address this, a DHI was proposed to maintain patient contact after RPM ends, as patient involvement with the clinical team decreases post-RPM.

3.2. Objectives for a solution

3.2.1. Design search

The literature review results are detailed in Appendix A, revealing the necessity of a multidisciplinary team. These studies used rule-based models to deliver messages tailored to selected target behaviors. Three studies employed theoretical models to design text messages [26–28]. Inspired by these, we described the system's complexity using the Behavior Change Wheel [29]. This involved defining the target behavior, identifying barriers and facilitators to self-management through qualitative methods, and mapping these to the COM-B model to understand necessary changes. Behavior change techniques with evidence of leading to the target behavior were used to develop the content sent to patients [30].

3.2.2. Prototyping session

A prototyping session with the members of the clinical team (n = 3) involved in the management of patients after hospital discharge. The results are shown in Table 2.

3.3. Design and development

3.3.1. Platform

The platform was designed in accordance with the specifications previously defined in Table 3.

The platform was developed using a modular architecture that divides the platform in different apps, to encapsulate all the details of a specific process, as described in Table 4.

3.3.2. Content design

A rule-based conversational agent was deployed given the results from the conversation flow design session. An example of interaction is shown in Table 5.

Table 2
The artifact categories discussed during the prototyping session, along with the rationale behind each decision.

Category	Decision	Reason
Artifact type	Conversational Agent	Clinical team experience, and prior literature.
Target behavior	Healthy Diet, and Physical activity	Relationship with other modifiable risk factors; Perceived frequency by which each patient does these behaviors (Fig. 5); BMI (Table 1).
Communication Channel	SMS	Age, education, and socioeconomic status (Table 1).
Dose	Four times per week, and customizable by the patient.	Clinical team experience, and prior literature.
Timing	Customizable by the patient.	Clinical team experience, and prior literature.

Table 3
Artifact specifications, the evaluation strategies to test the MVP.

Specifications	Platform to convey messages personalized to patient's needs; A conversational agent fully automated should improve patients' ability to self-manage their condition; It should be modular to be able to integrate new behavioral models, and be able to scale up; It should include mechanisms to handle errors; Healthcare providers should be able to validate the content sent to patients; The messages should mimic the way nurses provide instruction to patients and be encouraging, positive, friendly, polite and lighthearted; Personalization features: content, communication channel, and dose.
Demonstration and Evaluation	Feasibility study; Pilot Study; Focus Group.

Table 4
Apps within the platform.

Core app	The Core app serves as the backbone of the Django project, encompassing settings, WebSocket implementation, and server-related configurations.
Communication system	This app orchestrates the tasks essential for managing communication between the system and end users. Employing an asynchronous approach that enables the system to efficiently register new tasks in queues as incoming requests arrive, ensuring seamless responsiveness.
Users	This app is responsible for creating, updating, validating, and displaying user profiles of the user on the interface. It serves as the central hub for managing user-related functionalities.
Content	The app facilitates the creation, listing, and updating of messages, as well as the importation of new messages and questions from an Excel file. Additionally, it provides the capability to associate behavior change techniques, and mechanisms of action. Users can also impose restrictions, such as excluding certain messages for patients taking anticoagulant medication.
Plans	This app offers insights into a patient's plan, enabling the updating of the plan itself, listing associated messages and questionnaires, and adjusting the date and frequency of message dispatch. Plans play a crucial role in organizing and managing the flow of communication within the rule-based conversational agent.
Patients	The Patients module facilitates the creation of new patient profiles either through file import or individual entry. It provides functionality to list patients and update their profiles, ensuring comprehensive management and accessibility of patient information.

3.4. Demonstration and evaluation

3.4.1. Feasibility study

Based on the SUS results, 71 % of participants found the platform suitable for its intended purpose. Additionally, all participants unanimously agreed that the content provided, and the number of messages sent were apt for its intended use.

3.4.2. Pilot study

Acceptability was 100 %, and retention was 75 %. One participant provided an incorrect phone number, resulting in their dropout. Three participants were above 55 years old (P1, P2, and P4), and one participant was below 40 years old (P3). The conversation flow is shown in Table 5.

The system was able to send all text messages to the participants. However, there were issues while sending the weekly questionnaires to the participants. Issues with the bidirectionality of the system were later fixed, and the third participant (P3) received and answered all the questions. This technical issue made us not able to measure engagement

Table 5
Example of text messages to be sent to patients who chose both physical activity and healthy diet as target behaviors.

Date	Message
12/06/2023	{Greetings} This week we'll talk about nutrition. Having a balanced diet is crucial for your health. It helps better control cholesterol, blood pressure, and blood sugar levels.
13/06/2023	{Greetings} Today we will review what a healthy diet is. A healthy diet includes plenty of vegetables and fruits, carbohydrates like whole-grain bread, pasta, and rice, lean or low-fat dairy, more fish than meat, water as needed, and contains little fat, salt, and sugar.
14/06/2023	{Greetings} Reducing salt in food helps keep blood pressure lower. Therefore, try using more spices and reducing the amount of salt you use when cooking.
19/06/2023	{Greetings} This week we'll talk about physical activity. Science shows that making changes to your diet and physical exercise is more effective than prescribed medications in preventing diabetes!
20/06/2023	{Greetings} Reversing old habits can be difficult, so start with small things: try taking the stairs instead of the elevator, park the car further from home, or walk!
21/06/2023	{Greetings} Consider reducing your sitting time by taking a walk after dinner – just 15 min of walking can make a difference!
26/06/2023	{Greetings} This week we'll talk about nutrition. Getting rid of old habits is challenging, so start with simple things: try eating fruits when you crave something sweet or drink water instead of juice.
27/06/2023	{Greetings} Some people find it helpful to check the sugar content on food labels when shopping. Think of ways to cut some sugars from your diet.
28/06/2023	{Greetings} Today, we leave you with an idea: if you track your food choices, you'll realize which foods can be exchanged to improve your diet.
03/07/2023	{Greetings} We're in the last week. You'll miss these messages! Let's focus on physical activity. If you exercise daily for 30 to 60 min, you may be improving your health, reducing the risk of hypertension, high cholesterol, and obesity!
04/07/2023	{Greetings} Today, remember that physical activity not only helps you live longer but also helps prolong your health, mental agility, and independence for a better quality of life.
05/07/2023	{Greetings} It's the small steps that lead to big changes in your health. Every step you take this week will contribute to your health.

in the first two participants.

All participants found the system suitable for their daily life due to its simplicity. P1, P2, and P3 found the content provided useful. The participants tried to exchange messages with the CA when there were no questions being posed by the CA, which reveals the need to ensure bi-directional communication in the next iteration.

P1 did not find that these messages were effective, as the diet and physical activity were already good. P2 found the messages effective in facilitating improved physical activity levels, noting that they served as helpful reminders to engage in physical activity. P3 found that these messages were encouraging and helped improve both target behaviors. P1 did not feel the need to have a smart system that would answer any questions. P2 and P3 expressed the belief that a more intelligent system would be beneficial. P2 articulated the desire for personalized recipe recommendations based on specific ingredients, while P3 proposed leveraging AI technology for this purpose.

3.4.3. Focus group

The clinical team expressed appreciation for the solution, recognizing its potential as a valuable tool to reinforce patient recommendations.

Following the UTAUT, the latent variables that determine user acceptance and use behavior are performance expectancy, effort expectancy, social influence, facilitating conditions, and behavior intention.

The solution aims to improve performance expectancy by delivering essential information to address risk factors and empower patients, especially those not adhering to recommendations by the third month. The effort required from the clinical team would be minimal given system's automation. The head doctor and head nurse represent a positive social influence as they have greatly influenced the adoption of digital solutions, embedding this practice into the cardiothoracic surgery service culture. Key facilitating conditions included the need for automation given patient's clinical history and nurse's intervention, and integration with the hospital's current platforms to enhance care efficiency.

In regard to behavior intention, the clinical team believes this solution would be beneficial after the RPM period for patients not following recommendations. While not all patients will use it, those in need would benefit greatly.

The nursing team suggested integrating additional targeted behaviors (reducing alcohol consumption, wound care, and medication adherence), and emphasized the importance of personalization based on the patient's history and environment.

The team acknowledged the challenges of using AI for message generation due to the need for accurate information. They expressed openness to testing such a system if the hospital had the necessary infrastructure, preferring bi-directional communication as developed during the pilot. They favored receiving phone calls from patients for specific explanations over focusing on patient-sent messages.

Lastly, the team suggested conducting a new study with more patients to gather evidence on the solution's effectiveness.

4. Discussion and conclusion

This study led to the creation of a platform to send messages to help patients enhance their self-management skills after cardiothoracic surgery. The use of the DSR methodology guided the development of a system that incorporates input from all stakeholders, may ensure the solution's sustainability and enabling rapid iterations over the product [7]. By doing so, we intended to create a solution able to be generalized to other contexts. To the best of our knowledge, no such solution has been created within the context of cardiothoracic surgery. This article's primary contributions lie in detailing the process behind the development of a platform designed for communicating post-surgery messages to patients. It also encompasses its thorough evaluation and identified factors crucial for enhancing its impact, including clinical team adherence and patient engagement. Minor contributions include design strategies that have been used to develop such systems, the factors to be used for personalization and the content that is perceived to be more relevant for the clinical team.

Modifiable risk factors were unanimously seen as the most important factors to ensure secondary prevention to the clinical team. In future studies, modifiable risk factors should be assessed to evaluate the solution's effectiveness.

Personalization features found at first resided on the selection of target behaviors that could improve patient's self-management skills, its barriers, and facilitators to performing a specific behavior, and user experience factors, such as, the communication channel, the timing, and the dose. After the pilot, the clinical team found that in addition to physical activity and a healthy diet, the system should be able to focus also in reducing alcohol consumption, proper surgical wound care, and medication adherence, as these are interventions they frequently encounter. Additionally, the clinical history of the patient should modulate the content provided to ensure patient's safety and engagement, as discussed in the focus group.

Although this solution was specifically tested in the context of cardiothoracic surgery, we believe it has the potential to be applied in other clinical settings. Its design allows for generalization to a variety of diseases and acute conditions, making it adaptable to different healthcare contexts.

The pilot study had limitations, including a small sample size of 4 participants and a short intervention period of only one month. Future studies should involve a more representative sample and include multi-site trials to ensure the solution's relevance for a broader population. To better understand factors such as non-adherence, usage patterns, and participant needs, future research should stratify participants to define patient profiles and explore tailored engagement strategies. Long-term follow-up studies are needed to assess the sustained impact of the intervention. Additionally, insights from the pilot study revealed that some patients would like the CA to respond to more complex inquiries, suggesting a need for bi-directional communication. Various methods, such as rule-based models, retrieval-based models, or AI-based models, can be employed to achieve this [24]. Implementing such a system in the future, may improve user engagement, and the effectiveness of the intervention.

The focus group revealed that there is the need to ensure the interoperability of this solution with the other service systems and platforms. However, due to the heterogeneity of different information systems, it is difficult ensuring the synchronous communication, exchange and use of information effectively and efficiently [31]. Future solutions must address various interoperability levels using the standards that have been developed by entities as Health Level Seven [31].

LLMs are showcasing the possibility to approach human proficiency [32]. However, LLMs are known to occasionally generate unreliable information, in a phenomenon known as hallucination [33]. To mitigate this risk, different methods are required to ensure the reliability of the information provided to the end-user [34]. Retrieval-augmented generative models can improve content accuracy by grounding responses in verified information[35]. Hallucination detectors have also been used to mitigate these issues [36], although these models have been having issues while running in real time, researchers are focused on improving their latency [37].

Healthcare providers need reliable systems that can be trusted. Therefore, future research may focus on the use of such a model to attend to user's needs and ensure the efficacy, interpretability, applicability, robustness, scalability, privacy, fairness, and justice [38]. Specifically, human-in-the-loop approaches—where clinicians validate or supervise AI-generated content—could further enhance safety and trustworthiness. Moreover, ongoing model evaluation and updates will be essential to continuously improve content safety as AI systems evolve.

In conclusion, this study marks a significant step towards enhancing post-cardiothoracic surgery care through conversational agents. The integration of conversational agents within the healthcare landscape presents an avenue for personalized, effective, and accessible patient care. Through collaboration, innovation, and a commitment to patient-centric solutions, we strive towards a future where technology seamlessly augments and enhances the healthcare experience. We believe that in the long term, this solution could improve health outcomes through increased adherence to treatment plans, better symptom management, and an increased well-being to patients. Healthcare providers may benefit from such a solution to automate routine communication tasks. Moreover, integrating the system with existing platforms, as highlighted by healthcare professionals in the focus group, could streamline workflows, improve information access, and enhance coordination across various healthcare services. Long-term, this could lead to more efficient healthcare delivery, reduces, administrative burdens, and improved care coordination, ultimately enhancing the quality of patient care and system sustainability.

5. Authors' contributions

AM, AL, LVL, ILN found the methodologies to be used within each DSR iteration.

AM, AL, LVL, ILN did the literature review.

AM wrote the interview guide for the in-depth interview with the clinical team. AL, LVL, and IN validated the guide. RS ensured that the questions were well formulated. Data analysis was performed by AM, and validated by AL, LVL, and ILN.

AM wrote the interview guide for the interview with the patients. AL, VLL, IN validated the guide. RS ensure that the questions were well formulated. Data analysis was performed by AM and validated by AL, LL, and IN.

AM, AL, LVL, ILN, CV, HS, PC defined the objectives for the solution.

AM developed the platform and gathered content from prior studies.

RS, CV and HS validated the content.

CV and HS recruited the patients for the pilot study.

AM interviewed the patients after the pilot study.

AM and APG prepared and conducted the focus group with the nursing team.

AL, LVL, ILN, CV, HS, PC, APG reviewed the full manuscript.

6. Ethical declarations

The clinical study was approved by the Ethical Committee of Centro Hospitalar Lisboa Central under study registration number 1187/2022. The ethical review process involved a comprehensive evaluation of the study's potential risks and benefits, ensuring compliance with local and international ethical standards. The Committee confirmed that the study met all ethical requirements for participant safety, welfare, and data protection.

All participants in the pilot study provided written informed consent prior to their involvement in the study. Consent forms were made available in the participants' native language, and individuals were given the opportunity to ask questions and seek clarification. Participation in the study was entirely voluntary, and participants were informed that they could withdraw at any time without any consequence.

To minimize potential risks, a rule-based conversational agent was chosen to ensure the delivery of reliable and consistent content to patients. If a patient sent a message outside of the predefined message plan, the contact number for the cardiothoracic surgery department was automatically provided to ensure prompt and accurate assistance.

The study strictly adhered to the principles of the Declaration of Helsinki, ensuring that participants' rights, safety, and dignity were protected throughout the research process.

In terms of data privacy, all personal data was handled in accordance with the European General Data Protection Regulation (GDPR). Participant data was anonymized at the point of collection, and all identifying information was removed to protect confidentiality. Data was securely stored on encrypted servers with access restricted to authorized personnel only. Participants were informed of their data rights under GDPR, including the right to access, rectify, or request the deletion of their data.

The anonymized data will be retained for five years following the study's conclusion, in line with institutional data management policies. Any future use of the data will require additional approval from the Ethical Committee to ensure ongoing compliance with ethical standards.

7. Summary Table

- Effective self-management is essential for long-term survival and reduced need for further interventions.
- There is lack of evidence on the use of conversational agents can enhance the recovery of patients after cardiothoracic surgery by improving their self-management skills.
- Employing iterative and co-design methods adds value to the solution, ensuring its sustainability.
- This study accessed the usability, acceptability, retention, and perceived effectiveness of the use of a conversational agent in patients after cardiothoracic surgery. Additionally, the conditions relevant to ensure the clinical team acceptability were assessed using qualitative methods.
- It is crucial to ensure that digital solutions are seamlessly integrated into the workflows of healthcare professionals.
- Leveraging generative artificial intelligence can significantly enhance patient engagement; however, it is crucial to ensure the technology's reliability and trustworthiness.

CRediT authorship contribution statement

Ana Martins: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Luís Velez Lapão:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Isabel L. Nunes:** Writing – review & editing, Validation, Methodology, Conceptualization. **Ana Paula Giordano:** Writing – review & editing, Validation, Methodology. **Helena Semedo:** Writing – review & editing, Validation, Conceptualization. **Clara Vital:** Writing – review & editing, Validation, Conceptualization. **Raquel Silva:** Writing – review & editing, Validation, Methodology. **Pedro Coelho:** Writing – review & editing, Project administration, Conceptualization. **Ana Londral:** Writing – review & editing, Validation, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to acknowledge National Foundation of Science and Technology for funding this work under the projects DSAIPA/AI/0094/2020, UIDB/00667/2020 (UNIDEMI), and Lisboa-05–3559-FSE-3, and the research grant 2023.02916.BDANA. A special thanks to all the healthcare professionals from the Cardiothoracic Surgery Department of Santa Marta Hospital (Lisbon) for their invaluable contribution to this study. Additionally, the authors acknowledge the valuable contributions of the reviewers to the improvement of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2024.105640.

References

[1] A. Timmis, P. Vardas, N. Townsend, A. Torbica, H. Katus, D. De Smedt, et al., European Society of Cardiology: cardiovascular disease statistics 2021, Eur. Heart J. 43 (2022) 716–799, https://doi.org/10.1093/eurheartj/ehab892.
[2] S. Yusuf, P. Joseph, S. Rangarajan, S. Islam, A. Mente, P. Hystad, et al., Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study, Lancet 395 (2020) 795–808, https://doi.org/10.1016/S0140-6736(19)32008-2.
[3] N. Curcio, M.M. Bennett, K.R. Hebeler, A.M. Warren, J.R. Edgerton, Quality of Life Is Improved 1 Year After Cardiac Surgery, Ann. Thorac. Surg. 111 (2021) 1954–1960, https://doi.org/10.1016/j.athoracsur.2020.07.063.
[4] Capewell S, Lloyd-Jones DM. Optimal Cardiovascular Prevention Strategies for the 21st Century. JAMA 2010;304. doi: 10.1001/jama.2010.1641.
[5] E. Murray, E.B. Hekler, G. Andersson, L.M. Collins, A. Doherty, C. Hollis, et al., Evaluating Digital Health Interventions, Am. J. Prev. Med. 51 (2016) 843–851, https://doi.org/10.1016/j.amepre.2016.06.008.
[6] A. Londral, S. Azevedo, P. Dias, C. Ramos, J. Santos, F. Martins, et al., Developing and validating high-value patient digital follow-up services: a pilot study in cardiac surgery, BMC Health Serv Res 22 (2022) 680, https://doi.org/10.1186/s12913-022-08073-4.
[7] Hevner, March, Park, Ram. Design Science in Information Systems Research. MIS Quarterly 2004;28:75. doi: 10.2307/25148625.
[8] K. Peffers, T. Tuunanen, M.A. Rothenberger, S. Chatterjee, A Design Science Research Methodology for Information Systems Research, J. Manag. Inf. Syst. 24 (2007) 45–77, https://doi.org/10.2753/MIS0742-1222240302.
[9] Guthrie WF. NIST/SEMATECH e-Handbook of Statistical Methods (NIST Handbook 151) 2020. doi: 10.18434/M32189.
[10] Rossum G van, Drake FL. The Python language reference. Release 3.0.1 [Repr.]. Hampton, NH: Python Software Foundation; 2010.
[11] Adams WC. Conducting Semi-Structured Interviews. In: Newcomer KE, Hatry HP, Wholey JS, editors. Handbook of Practical Program Evaluation. 1st ed., Wiley; 2015, p. 492–505. doi: 10.1002/9781119171386.ch19.
[12] I. Masic, J. Alajbegovic, The significance of the psychosocial factors influence in pathogenesis of cardiovascular disease, Int J Prev Med 4 (2013) 1323–1330.
[13] L. Yang, A.D. Kaye, A.G. Venakatesh, M.S. Green, C.D. Asgarian, M.M. Luedi, et al., Enhanced Recovery after Cardiac Surgery: An Update on Clinical Implications, Int. Anesthesiol. Clin. 55 (2017) 148–162, https://doi.org/10.1097/AIA.0000000000000168.
[14] Jenkins CD, Stanton BA, Jono RT. Quantifying and predicting recovery after heart surgery.: Psychosomatic Medicine 1994;56:203–12. doi: 10.1097/00006842-199405000-00005.
[15] C.M. DuBois, O.V. Lopez, E.E. Beale, B.C. Healy, J.K. Boehm, J.C. Huffman, Relationships between positive psychological constructs and health outcomes in patients with cardiovascular disease: A systematic review, Int. J. Cardiol. 195 (2015) 265–280, https://doi.org/10.1016/j.ijcard.2015.05.121.
[16] LimeSurvey Development Team. LimeSurvey - The free and open source survey software tool! LimeSurvey Development Team; 2012.
[17] J. Saldaña, The coding manual for qualitative researchers, 2nd ed., SAGE, Los Angeles, 2013.
[18] Martins A, L. Nunes I, Velez Lapão L, Londral A. Conversational Agent Automation for patients after cardiothoracic surgery based on clinical team experience and behavior change theory, 2023. doi: 10.54941/ahfe1003597.
[19] Django Software Foundation. Django 2019.
[20] M. Sadek, R.A. Calvo, C. Mougenot, Co-designing conversational agents: A comprehensive review and recommendations for best practices, Des. Stud. 89 (2023) 101230, https://doi.org/10.1016/j.destud.2023.101230.
[21] M. MacPherson, K. Cranston, C. Johnston, S. Locke, M.E. Jung, Evaluation and Refinement of a Bank of SMS Text Messages to Promote Behavior Change Adherence Following a Diabetes Prevention Program: Survey Study, JMIR Form Res 5 (2021) e28163.
[22] Venkatesh, Morris, Davis, Davis. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27:425. doi: 10.2307/30036540.
[23] Martins A, Nunes I, Lapão L, Londral A. Designing a Digital Personal Coach to Promote a Healthy Diet and Physical Activity Among Patients After Cardiothoracic Surgery: Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies, Lisbon, Portugal: SCITEPRESS - Science and Technology Publications; 2023, p. 595–602. doi: 10.5220/0011776800003414.
[24] Martins A, Londral A, L. Nunes I, V. Lapão L. Unlocking human-like conversations: Scoping review of automation techniques for personalized healthcare interventions using conversational agents. International Journal of Medical Informatics 2024; 185:105385. doi: 10.1016/j.ijmedinf.2024.105385.
[25] Education System levels in Portugal | EURAXESS n.d. https://www.euraxess.pt/portugal/information-assistance/family-related-issues/education-system-levels-portugal (accessed December 27, 2023).
[26] L. Mayberry, C. Berg, R. Greevy, L. Nelson, E. Bergner, K. Wallston, et al., Mixed-Methods Randomized Evaluation of FAMS: a Mobile Phone-Delivered Intervention to Improve Family/Friend Involvement in Adults' Type 2 Diabetes Self-Care, Ann. Behav. Med. 55 (2021) 165–178, https://doi.org/10.1093/abm/kaaa041.
[27] J. Kelly, M. Conley, T. Hoffmann, J. Craig, A. Tong, D. Reidlinger, et al., A Coaching Program to Improve Dietary Intake of Patients with CKD: ENTICE-CKD, Clin. J. Am. Soc. Nephrol. 15 (2020) 330–340, https://doi.org/10.2215/CJN.12341019.
[28] Cadilhac D, Andrew N, Busingye D, Cameron J, Thrift A, Purvis T, et al. Pilot randomised clinical trial of an eHealth, self-management support intervention (iVERVE) for stroke: feasibility assessment in survivors 12–24 months post-event. Pilot and Feasibility Studies 2020;6. doi: 10.1186/s40814-020-00706-x.
[29] S. Michie, M.M. van Stralen, R. West, The behaviour change wheel: A new method for characterising and designing behaviour change interventions, Implementation Sci 6 (2011) 42, https://doi.org/10.1186/1748-5908-6-42.
[30] M.P. Guerreiro, J. Strawbridge, A.M. Cavaco, I.B. Félix, M.M. Marques, C. Cadogan, Development of a European competency framework for health and other professionals to support behaviour change in persons self-managing chronic disease, BMC Med Educ 21 (2021) 287, https://doi.org/10.1186/s12909-021-02720-w.
[31] Interoperability in Healthcare | HIMSS n.d. https://www.himss.org/resources/interoperability-healthcare (accessed September 5, 2024).
[32] E.P.H. Choi, J.J. Lee, M.-H. Ho, J.Y.Y. Kwok, K.Y.W. Lok, Chatting or cheating? The impacts of ChatGPT and other artificial intelligence language models on nurse education, Nurse Educ. Today 125 (2023) 105796, https://doi.org/10.1016/j.nedt.2023.105796.
[33] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, et al., Survey of Hallucination in Natural Language Generation (2022), https://doi.org/10.48550/ARXIV.2202.03629.
[34] Weidinger L, Uesato J, Rauh M, Griffin C, Huang P-S, Mellor J, et al. Taxonomy of Risks posed by Language Models. 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul Republic of Korea: ACM; 2022, p. 214–29. doi: 10.1145/3531146.3533088.
[35] A. Haleem, M. Javaid, R.P. Singh, An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges, BenchCouncil Transactions on Benchmarks, Standards and Evaluations 2 (2022) 100089, https://doi.org/10.1016/j.tbench.2023.100089.
[36] Zha Y, Yang Y, Li R, Hu Z. AlignScore: Evaluating Factual Consistency with a Unified Alignment Function 2023. doi: 10.48550/ARXIV.2305.16739.
[37] M. Hu, R. Xu, D. Lei, Y. Li, M. Wang, E. Ching, et al., SLM Meets LLM: Balancing Latency, Interpretability and Consistency in Hallucination Detection (2024), https://doi.org/10.48550/ARXIV.2408.12748.
[38] Ma Z, Ethayarajh K, Thrush T, Jain S, Wu L, Jia R, et al. Dynaboard: An Evaluation-As-A-Service Platform for Holistic Next-Generation Benchmarking 2021. doi: 10.48550/ARXIV.2106.06052.