Investigating Motivations and Patient Profiles for Personalization of Health Applications for Behaviour Change

Peer-reviewed author version

BONNEUX, Cindel; DENDALE, Paul & CONINX, Karin (2022) Investigating Motivations and Patient Profiles for Personalization of Health Applications for Behaviour Change. In: ACM - Association for Computing Machinery (Ed.). PETRA '22: Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments, ACM - Association for Computing Machinery, p. 146 -154.

DOI: 10.1145/3529190.3529201
Handle: http://hdl.handle.net/1942/39214

Investigating Motivations and Patient Profiles for Personalization of Health Applications for Behaviour Change

Cindel Bonneux
UHasselt, Faculty of Sciences, HCI and eHealth
Diepenbeek, Belgium
cindel.bonneux@uhasselt.be

Paul Dendale
UHasselt, Faculty of Medicine and Life Sciences
Diepenbeek, Belgium
paul.dendale@uhasselt.be

Karin Coninx
UHasselt, Faculty of Sciences, HCI and eHealth
Diepenbeek, Belgium
karin.coninx@uhasselt.be

ABSTRACT

Personalization is a key aspect when developing applications targeting health behaviour change. However, the use of personalized mobile interventions for lifestyle behaviour is still in its infancy. Based on our former research on mobile applications to support cardiac patients in health behaviour change, we identified four key motivations to enhance the personalization offered in applications targeting health behaviour change. In this paper, we propose a mixed-methods approach, using both qualitative and quantitative data collected in prior studies, to apply personalization in the design of health applications. Our approach consists of five steps: 1) collecting data for personalization, 2) detecting patient profiles using clustering methods, 3) understanding patient profiles using a graphical representation, 4) describing patient profiles using personas, and 5) personalizing a health application according to patient profiles. One of the major strengths of our approach is that it combines established HCI techniques such as personas and data visualization techniques with methods from big data analytics and artificial intelligence to identify ways to personalize health applications. We conclude by presenting future directions to apply personalization in the domain of health technologies.

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**; • **Applied computing →** *Health informatics*; • **Information systems → Personalization**.

KEYWORDS

personalization, behaviour change, eHealth, user-centered design, approach

ACM Reference Format:
Cindel Bonneux, Paul Dendale, and Karin Coninx. 2022. Investigating Motivations and Patient Profiles for Personalization of Health Applications for Behaviour Change. In *The15th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '22), June 29-July 1, 2022, Corfu, Greece.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3529190.3529201

ACM ISBN 978-1-4503-9631-8/22/06…$15.00
https://doi.org/10.1145/3529190.3529201

1 INTRODUCTION

Cardiovascular diseases (CVDs) are world's leading cause of death, an estimated 17.9 million people died from CVDs in 2019 [23]. After a cardiac incident, patients typically enrol in a cardiac rehabilitation (CR) program to foster recovery and reduce their risk of recurrent events. A comprehensive cardiac rehabilitation program is composed of several key components aiming at a healthy lifestyle: parameter monitoring, education, medication, physical activity, nutrition, smoking cessation, and stress management [2]. Depending on their risk profile, patients need to make different modifications to their lifestyle. Some patients already have quite a healthy lifestyle and need to make only a few changes (e.g. make some small modifications to their eating patterns, such as eating less salt), whereas others have a very unhealthy lifestyle and need to drastically change their life (e.g. quit smoking, start exercising and eat more healthy). Therefore, personalization plays a key role in adapting the cardiac rehabilitation program to the patient's risk profile, and can optimally be supported by personalization techniques in the mHealth applications that come into play during the rehabilitation trajectory. In addition, it is important that patients maintain this healthy lifestyle for their entire life.

1.1 Personalization

In the context of IT systems, personalization can be defined as a "process that changes the functionality, interface, information access and content, or distinctiveness of a system to increase its personal relevance to an individual or a category of individuals" [4]. Personalization has been applied in numerous applications, including games for education [20], e-commerce [25], recommender systems [5] and health applications [9].

According to the European Society of Cardiology, there is an urging need to transform the current stratified practice of cardiovascular disease management into a better personalized cardiovascular medicine, fitting within the broader context of global patient care [8]. However, the review of Tong et al. [19] demonstrated that the use of personalized mobile interventions for lifestyle behaviour is still in its infancy. Existing applications mostly personalize content and rarely personalize other features, such as intervention timing, dosage, or delivery. Furthermore, they demonstrated that effectiveness was higher for interventions using system-captured data for personalization than for interventions using user-reported data, highlighting interesting directions for future research. In this paper, we identify key motivations why personalization is needed for eHealth applications targeting behaviour change.

1.2 HeartHab

In this paper, we extend the analysis of some of our former studies in cardiac rehabilitation from the perspective of individual application usage. Based on this analysis, we identify key motivations to personalize eHealth applications targeting behaviour change. Our studies used two variations of the HeartHab application. HeartHab [18] is a mobile app that supports cardiac patients in acquiring a healthy lifestyle by improving their medication adherence and increasing their physical activity. The HeartHab application is composed of five modules: medication adherence, physical activity, parameter monitoring, education, and symptom reporting. In the context of this paper, we define a module as a coherent set of features that relates to one of the rehabilitation components such as physical activity, medication, education, etc. Dependent on the module, the focus is on supporting the patient in changing his/her behaviour (behavioural module) or assisting the patient in becoming the manager of his/her disease (supporting module). The two behavioural modules of HeartHab are medication adherence and physical activity, whereas the three supporting modules are parameter monitoring, education, and symptom reporting.

All behavioural and supporting modules of HeartHab are available at all times. Within each module, there is personalization, such as personalized exercise targets, personalized feedback about parameter values, and tailored motivational messages. However, this personalization is rather on the level of feature mechanics (e.g. goal-setting and target ranges) and content (e.g. motivational messages), as opposed to full-blown personalization of features based on patient preferences, responsiveness to the intervention, and application usage. HeartHab was evaluated in a crossover trial with 32 coronary artery disease patients. All patients were randomized in a 1:1 ratio to telerehabilitation with the HeartHab application or usual care. All patients used the HeartHab mobile app for 2 months. Two patients met the exclusion criteria after randomization and two others withdrew from the study due to a medical reason [17].

Based on the results of the crossover trial, an evolutionary version of HeartHab was developed. This version of HeartHab evolves over time, as the patient progresses in the rehabilitation trajectory. The behavioural modules medication adherence and physical activity are available for a fixed period of time, whereas the supporting modules parameter monitoring and education are available to the patient at all times. Furthermore, the educational content provided in the education module updates over time based on the patient's needs (from the patient's and caregivers' point of view) and preferences [16]. The supporting module to report symptoms was omitted in the evolutionary version of HeartHab, due to limited usage in the crossover trial and no active follow-up possibility by caregivers. The evolutionary version of HeartHab was evaluated in a small field study with 5 cardiac patients transitioning from rehabilitation in the rehabilitation center to home-based rehabilitation with remote monitoring. In the first week of the study, patients could use the parameter monitoring, education and medication adherence module. At the start of week 2, the physical activity module was added to the application, to get the full HeartHab application. Next, the two behavioural modules were again gradually removed. The medication module was removed after week 2 and the physical activity module after week 4. Two patients quit the study prematurely due to developing co-morbidities or other personal reasons.

In the Section 2, we look at the results of both studies using the HeartHab application. We refer to the 2 studies as respectively the "crossover trial" and the "evolutionary study".

1.3 Patient profiles in user-centered design

User-centered design approaches are extensively employed to design health applications for diverse patient populations. A first step in a multidisciplinary user-centered software engineering process is collecting and understanding requirements and user needs. As a result of the user needs analysis, narrative artifacts such as personas and scenarios can be created [6]. Personas represent hypothetical archetypes of target users [12], whereas scenarios may include personas and describe how the future system will be used [3]. In this paper, we are interested in the development of personas as a technique to identify directions for applying personalization in the design of applications supporting health behaviour change.

Mulder and Yaar [11] define three approaches to create personas: 1) qualitative personas, 2) qualitative personas with quantitative validation, and 3) quantitative (or data-driven) personas. However, other researchers also refer to hybrid personas that use mixed methods [14]. Mixed methods, using both qualitative and quantitative data, is even advised for creating personas [13]. Therefore, we focus in this paper on a mixed-method approach for persona creation. However, in essence, all methods for persona creation (qualitative, quantitative and mixed-methods) are composed of four key steps: (a) data collection, (b) segmentation and grouping, (c) analysis of the qualitative and/or quantitative data, and (d) creating persona profiles to present the user segments and their attributes as hypothetical archetypes of users [14, 22, 27].

The review of Alsaadi et al. [1] demonstrated that personas are adopted in different healthcare contexts. Data-driven, qualitative and mixed approaches to persona creation have been utilized. Data-driven approaches were most prevalent in a medical context, whereas a mixed approach was only applied in one of the included studies. Qualitative and mixed persona development suffered from one major limitation, i.e. limited data size to create personas. An example of data-driven persona creation is the study of Holden et al. [7]. In their research, they developed biopsychosocial personas from a study of older adults with heart failure. They used standardized surveys and medical record abstraction as input for hierarchical cluster analysis. For each of the resulting clusters, a persona was created.

Quantitative persona creation should go beyond mere persona creation by targeting real use cases. In a health context, tailored medical interventions can be designed for different subpopulations represented by personas and evaluated to see how health outcomes develop over time [14]. An example is the work of Vosbergen et al. [21] regarding tailoring educational messages to patient preferences. Based on data collected in an online survey, they performed k-means clustering [10]. For each resulting cluster, a persona was created and rules to tailor health education were defined. The resulting personas were later presented to patients, letting them choose with which persona they could identify best. Based on the chosen persona, tailored education was provided.

1.4 Our contribution

The main contributions of this paper are twofold. First, we analyze data of our own prior studies on a persuasive health application to get insight into why personalization is very relevant for health applications supporting behaviour change. From this analysis, we identify four key motivations to personalize applications targeting health behaviour change. Next, we focus on how such a personalized health application can be designed by using data of prior studies. Given the low application of mixed-methods approaches for persona creation in healthcare [1], we develop an approach that combines both qualitative and quantitative data to define personas representing different patient profiles. The main difference between our approach and existing approaches for persona creation (e.g. data-driven personas) is the addition of an extra step using a graphical representation of patient data. This graphical representation can be used for the creation of personas and as a next step, for the design of personalized health applications. We conclude by highlighting some directions for future research.

2 KEY MOTIVATIONS TO PERSONALIZE HEALTH APPLICATIONS

In this section, we analyze data of our previous studies on the HeartHab application to grasp why personalization is essential for health applications targeting behaviour change. During both studies, logs of application usage were collected, detailing which screen was accessed by the patient and when it was accessed. We analyzed these logs to determine patients' needs for the different behavioural and supporting modules, and to get insights in patients' objective engagement with the mobile app. We detected usage sessions of modules by identifying subsequent accesses to screens of the same module. Every time there was a change in module, we considered this a new usage session. In addition, when there were at least 15 minutes between two navigation interactions, we considered this a new usage session due to the large time of inactivity. At the end of both studies, an individual, semi-structured interview was conducted with each study participant. We analyzed the interview data to explain patients' usage patterns patterns and pinpoint motivations for personalization.

2.1 Variation over time

When starting to work on a new behaviour change, we would expect that people need a lot of support. This support can be provided by a health application, but also human support from e.g. caregivers and family is important. At the start of a study, we would expect that patients use the mobile app multiple times a day, reflecting the novelty of the application but also the need for support in making the behaviour change. As patients are working longer on the behaviour change and gradually developing a new habit, we would expect that they become more independent and need less external support to sustain the behaviour change, which would result in lower application usage. For the supporting modules, we expect a rather similar usage pattern as for the behavioural modules, with the exception that symptom reporting and parameter monitoring would continue constantly, allowing patients to follow up on their behaviour change process. For education, we would expect that patients need more education in the beginning, when they start working on behaviour changes. Later on, they could go back to the educational content to rehearse important concepts or look something up.

In Fig. 1, the average usage of all modules of HeartHab over time during the crossover trial is depicted. Note that we left out day 1 of the study, because on this day there was a peak in usage, due to the application demonstration given by the researcher. We can see that over time, the usage of the health application decreases. The parameter monitoring module is used the most on average, followed by the medication adherence module, and the physical activity module. The symptom reporting and education module are not used extensively. For education, this could be devoted to the fact that the patients that participated in this trial already finished their rehabilitation quite some time ago (participation to the study was offered as boost therapy), which could mean that they already received the required information earlier, during their supervised rehabilitation. For example, PAT012 clearly stated in the semi-structured interview that he/she only watched one video because he/she already received all information from the hospital. For the symptom monitoring, the low usage could be devoted to the fact that symptoms reported in the mobile app were not monitored by the caregivers. In the interview, most patients indicated that they would find it useful to report symptoms if their caregivers would actively follow up the reported symptoms. However, some patients still preferred to have the opinion of the doctor regarding symptoms.

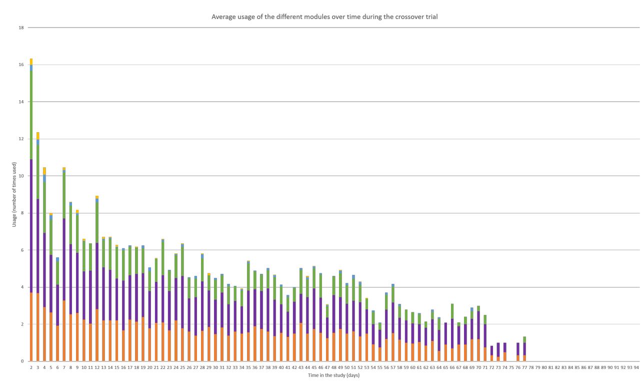

Figure 1: Average usage of the different modules over time during the crossover trial of HeartHab.

2.2 Differences between individual patients

Each patient has a unique profile, consisting of personal needs, preferences, motivations, goals, etc. When looking at a specific behavioural module, we would expect that some patients have a higher need for support than others. Based on interviews after the study, we conclude that there are multiple reasons for a detected higher usage, including needing more support for the behavioural goal, liking the features offered in the behavioural module, and enjoying the activities related to the behavioural goal (e.g. patients that like to do sports).

Fig. 2 illustrates the usage of the physical activity module of HeartHab by three different patients during the crossover trial. We chose these three patients, since they had varying usage patterns for the physical activity module. From these three patients, PAT018 used the physical activity the most, almost daily. PAT002 also used the physical activity module frequently, whereas PAT030 used the module only sporadically. For both PAT002 and PAT018, there is quite a constant usage of the physical activity module over time. This could be devoted to the fact that the patient did not change his/her behaviour yet, enjoyed exercising in general, or liked the support of the physical activity module. To be able to pinpoint this, we should look at the relation with the outcome values (in Section 2.4) and qualitative data collected in interviews and questionnaires.

For this particular example, we found a rationale for patients' usage patterns by analyzing the data of the interviews. PAT018 mentioned that the application encouraged him/her to increase his/her efforts for physical activity. In contrast, PAT002 mentioned that he/she was already physically active, so the application did not encourage him/her to be more physically active. Nevertheless, he/she still found the visualizations offered in the application motivating, indicated by the frequent usage of the physical activity module. Lastly, PAT030 had problems using technology in general and as a result did not use the application (including the physical activity module) a lot.

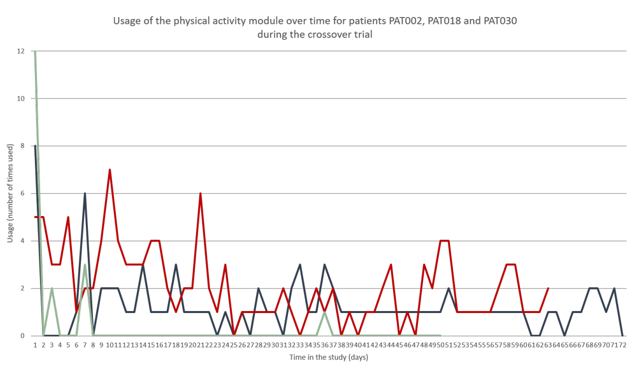

Figure 2: Usage of the physical activity module over time for patients PAT002, PAT018 and PAT030 during the crossover trial of HeartHab.

2.3 Differences between modules

Dependent on the module at hand, i.e. the target behaviour (for a behavioural module) or the provided support (for a supporting module), a patient has different needs for support. When looking at medication, some cardiac patients already have a good medication adherence and do not need frequent reminders and support in improving their medication adherence (e.g. PAT012, PAT017, and PAT018). However, others have a very hard time taking their medication regularly, requiring extensive support to work on improving their medication adherence (e.g. PAT015, PAT020, and PAT021). Each behaviour change has some unique characteristics, requiring different behaviour change techniques. In addition, some patients might have more trouble working on multiple behaviour changes simultaneously, whereas for others this is not a big challenge and they easily use the application for diverse behavioural goals. Also, patients' needs for a particular target behaviour change over time, e.g. based on changes in their risk profile or motivation.

During the crossover trial of HeartHab, all patients had all modules available at all times. In Fig. 3, the usage of the different modules over time for patient PAT007 is depicted for the crossover trial. We chose patient PAT007, because the usage frequency for the different modules was highly variant. Also, this patient indicated in the interview that he/she would definitely want to continue using the application after the trial, indicating that this patient is an adopter of our application. The most frequently used module for this patient was parameter monitoring, highlighting the patient's desire for parameter follow-up. However, the high number of usages of the module can be devoted to the parameter overview being the home screen of the application, but also the patient's perceived usefulness of the parameter overview. When looking at the behavioural modules, the medication adherence module was used more than the physical activity module. This could be devoted to the fact that the patient valued the medication reminders, that made sure he/she did not miss a single intake moment. Furthermore, taking medication is a daily activity for all cardiac patients, whereas not all patients succeed in incorporating daily exercise training in their life. The patient more frequently used the behavioural modules medication adherence and physical activity than the supporting modules education and symptom reporting. Especially the symptom reporting module was not used frequently, which is a general trend that we observed for all patients. This is a positive finding, since patients only had to use this module when they experienced symptoms e.g. pain on the chest.

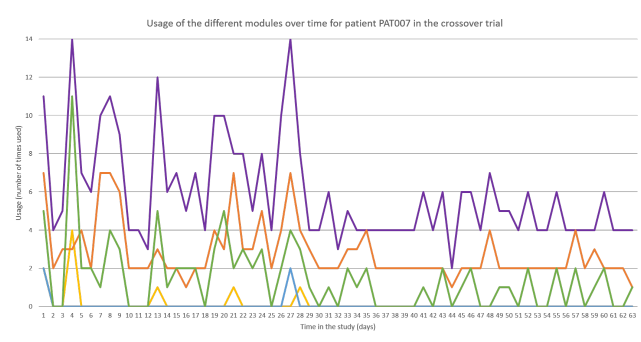

Figure 3: Usage of the different modules over time for patient PAT007 in the crossover trial of HeartHab.

In the evolutionary study of HeartHab, the module configuration was updated over time. In Fig. 4, the usage of the different modules during the evolutionary study is depicted for patient PAT042. During the first week, the patient had only access to the behavioural module medication adherence and the supporting modules parameter monitoring and education. During the second week, all behavioural and supporting modules were activated and used by the patient. However, the usage of the newly added behavioural module (i.e. physical activity) was less than the already available behavioural module (i.e. medication adherence). At the start of week 3, the medication adherence module was removed. We notice that starting from week 3, the patient uses the application less frequently. In the last two weeks of the study, the patient only had the two supporting modules (i.e. education and parameter monitoring). However, these were only used occasionally by the patient.

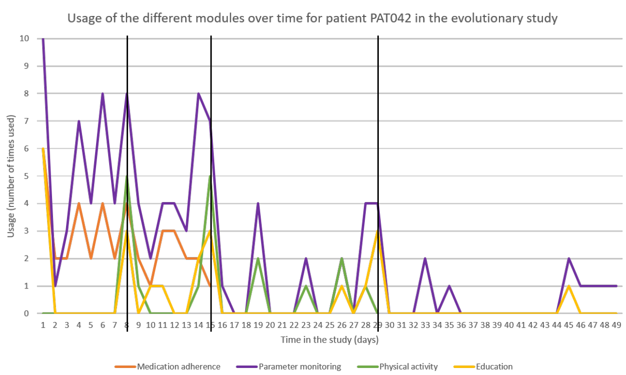

Figure 4: Usage of the different modules over time for patient PAT042 in the evolutionary study of HeartHab.

It is remarkable that when a new behavioural module was introduced (i.e. physical activity), the usage of the other behavioural module (i.e. medication adherence) did not drop noticeably, especially because the patient indicated in the interview that he/she found the registration of medication intake useless because he/she was already compliant. However, the patient would prefer to have all functionalities available at all times, which would positively influence his/her application usage since as there are more features he/she uses the application more often. We believe that the changes in module usage when the application is updated (e.g. adding/removing a module) are patient- and module-dependent.

Furthermore, it can be noticed that in the evolutionary study the education module was used more often than in the crossover trial. A possible explanation could be that the patients liked the fact that the education changed over time, accounting for their needs and preferences. Another reason could be the difference in study population, since in the evolutionary study patients that were finishing the rehabilitation program in the rehabilitation center were recruited. The analysis of the questionnaire data revealed that all patients felt better informed. Furthermore, for all three patients, the perceived usefulness of the tailored educational videos increased during the study. The continuous usage of the supporting modules (i.e. parameter monitoring and education) throughout the entire evolutionary study highlights patients' need for educational information and desire for follow-up of their medical values (e.g. weight).

2.4 Correlation with outcomes

As a last analysis, we look at the relation between the usage of a behavioural module and the related outcome parameter. We hypothesize that when patients have changed their behaviour and thus have good results for the related outcome parameter(s), they would use the related behavioural module less. However, looking at the data learns us that for some patients it is hard to see this relationship clearly, as for example for PAT001 in Fig. 5. However, there are also other possible reasons why the usage of a certain module could decrease or even drop to zero. E.g., a patient could become demotivated and stop working on the behaviour change or alternatively, the patient could lose interest in the application, or seek other support for the behaviour change. To be able to grasp why a patient uses a behavioural module less or even abandons the application, a correlation analysis of the behavioural module usage and the outcome parameter can be made. For this paper, we focused on a graphical approach for analyzing this relationship (Fig. 5), which is complementary to the more detailed mixed-methods analyses that was already performed [17].

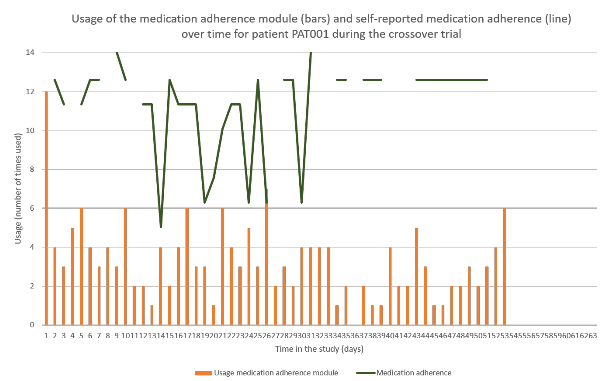

Figure 5: Usage of the medication adherence module (bars) and self-reported medication adherence (line) over time for patient PAT001 during the crossover trial of HeartHab.

Another reason to consider correlation analysis of usage data and outcome parameters is to detect if the behavioural module is effective at improving the patient's behaviour. When the behavioural module is effective, a high usage of the module should result in an improvement in the outcome parameter. However, when following this reasoning, additional data about other parallel interventions should be collected, so these can be taken into consideration. In Fig. 5, we see that the medication adherence of patient PAT001 is quite varying over time, despite almost daily usage of the medication module. The average self-reported medication adherence is quite good (81.28 percent). We can thus conclude that on most days, the medication module achieved its intended effects. We noticed that for this patient, almost every time that there is a low self-reported medication adherence, this was preceded or coincided with a lower usage of the medication module. Furthermore, we see that at day 54 the patient stopped using the medication module and thus did no longer report his/her medication adherence.

2.5 Summary

Based on the analysis of our former studies with both the static and the evolutionary version of HeartHab, we identified four key motivations to personalize health applications targeting behaviour change:

- Each patient has a unique profile, consisting of personal values, needs, goals, and preferences (Section 2.2);
- Patient needs evolve over time (Section 2.1);
- Patient needs are dependent on the module or behavioural goal (Section 2.3);
- Patient needs can be correlated with outcome measures to evaluate if the application has the intended effects (Section 2.4).

In our analysis, we mainly looked at moment-to-moment engagement with the application in terms of application usage. However, studies have drawn attention to issues with using only behavioural indicators to assess engagement [24]. E.g., the most commonly cited problem is lurking, i.e. when users are just looking at the screen without performing the expected actions. We recognize the need to combine objective measures of engagement with experiential engagement that can be assessed using questionnaires and interviews, which can give us a detailed insight if patients' engagement with the application resulted in the desired impact on their motivation and supported them in their behaviour change process.

Furthermore, qualitative research techniques are essential to provide an explanation for the application usage patterns. For example, when there is a drop in application usage for a certain module, this could be caused by a lack of motivation, disinterest in the application, or no need to work on the related behavioural goal. As we illustrated, qualitative techniques such as interviews or questionnaires can be used to gain a deeper insight and identify the reason for lower application usage or usage of a specific module. By combining these qualitative and quantitative methods, conclusions can be drawn about which personalization techniques work better for a specific patient in a particular context (as we did for the HeartHab crossover trial [17]).

3 PERSONALIZING A HEALTH APPLICATION BASED ON USAGE AND OUTCOME DATA

As we identified key motivations to personalize, a next step is to effectively personalize a health application. In this section, we present an approach to personalize a health application based on outcome and usage data that was collected in previous studies. Our proposed approach consists of the following five steps:

(1) Collecting data for personalization
(2) Detecting patient profiles using clustering methods
(3) Understanding patient profiles using a graphical representation
(4) Describing patient profiles using personas
(5) Personalizing a health application according to patient profiles

We envision our proposed approach to be used in the following three situations: 1) transform a one-size-fits-all health application into a personalized application based on data collected in previous studies using the application, 2) enhance the personalization of a health application based on data collected in previous studies using the application, and 3) design a new personalized application based on data collected from studies on similar applications. In our example of HeartHab, we are in the second situation, enhancing the personalization of the app based on data collected in prior studies. In the future, we see a fourth situation arising from dynamically adapting the application based on data that is collected during the study. Since this situation contains some characteristics of each of the three previously identified situations, similar techniques can be applied.

3.1 Collecting data for personalization

In studies on health applications various types of data can be collected, ranging from outcome data (e.g. weight) to data on the application usage (e.g. when does the patient use a certain feature). When aiming to design a new application or redesign an already existing application, data of relevant studies can be investigated as we do in research. However, we can go even one step further and use this data for the purpose of creating data-driven personas that drive the design of the application. For example, Zhang et al. [26] create behavioural personas from raw clicks gathered in telemetry data from actual application use. This study was one of the first to investigate the use of real application usage to create data-driven personas, instead of relying on data collected in surveys. In our method, we propose to not only rely on behavioural tracking data collected by actual application usage, but also outcome data, which is at least equally important in a health context, since using a health application does not necessarily lead to the intended effects (e.g. behaviour change, improved medical values, etc.). The strengths of combining quantitative and qualitative techniques in a mixed-methods approach has been demonstrated in the past. However, in the review of Salminen et al. [14] on quantitative persona creation only 7 out of the 49 included articles employed qualitative methods in both the initial data collection and the validation. In our approach, we suggest to augment the application data (usage and outcome) with data collected in surveys, interviews and user observations, resulting in a rich data set.

During both studies using HeartHab, we collected varying types of data. As discussed in Section 2, we collected usage data by logging when the patient accessed a certain screen. Qualitative data was collected in semi-structured interviews and custom-made questionnaires, whereas outcome data such as medical values (e.g. blood pressure) and behavioural data (e.g. sports activities) were collected through the app, questionnaires and the hospital information system. All collected information can be used to enhance the personalization of HeartHab.

3.2 Detecting patient profiles using clustering methods

After the data has been collected, the process of using the data to improve/design an application can be initiated. In general, we can consider two types of clustering: manual and automatic. In manual clustering, a person can detect clusters by analyzing the data manually by comparing patient characteristics. However, this is only feasible for small data sets. For larger data sets, automatic clustering using machine learning algorithms is preferred. In the "Cluster era" (2010-2013), a lot of research on clustering methods for data-driven personas has been performed. The most popular quantitative persona creation analytics methods are (in decreasing popularity) k-means clustering, hierarchical clustering, principal component analysis, latent semantic analysis, non-negative matrix factorization and lastly other newly introduced models [14]. We do

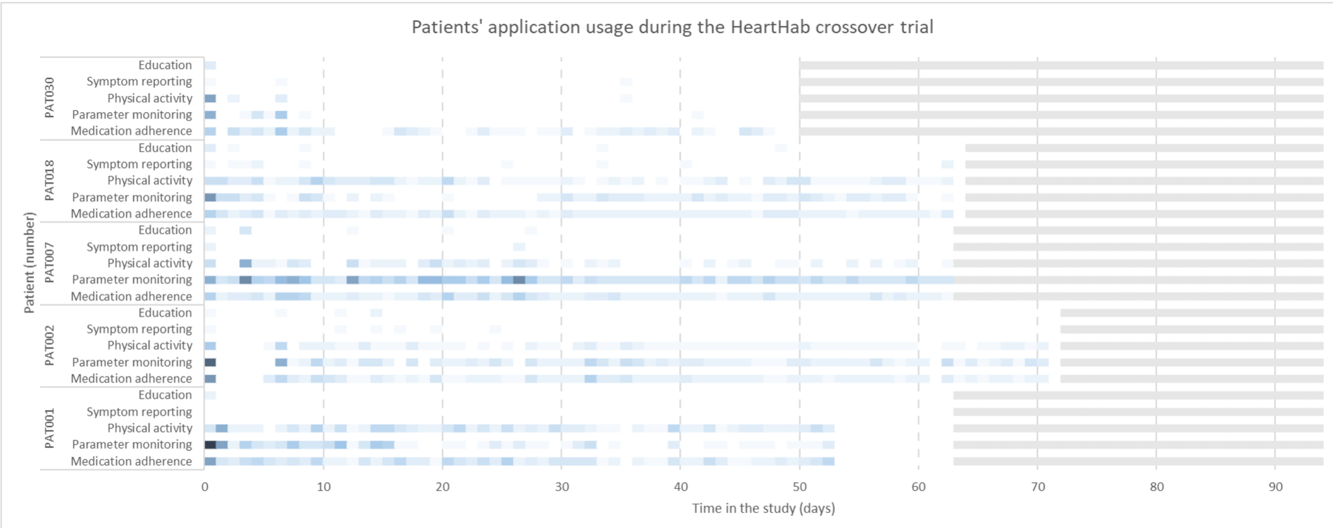

Figure 6: Graphical representation of patients PAT001, PAT002, PAT007, PAT018 and PAT030 in the HeartHab crossover trial.

not want to express a preference for any of these analytics methods for clustering the collected data. We suggest that researchers employ the most suitable technique for their context and the data at hand. After clustering methods have been applied, each resulting cluster can be transformed into one patient profile. In this paper, we consider a patient profile a set of patient characteristics (e.g. gender, usage patterns, medical condition) that is common for the patients allocated to a certain cluster, but significantly different from the other cluster(s).

Given the limited sample size and amount of data for the HeartHab studies, we performed manual clustering. We used a graphical approach to identify and select relevant clusters. The graphical representation introduced in Section 3.3 supported us in identifying clusters and patient profiles.

3.3 Understanding user profiles using a graphical representation

After detecting patient profiles, personas representing hypothetical archetypes of target users can be created [12]. However, creating these personas based on the identified clusters might be cumbersome and difficult depending on the number of resulting clusters and the magnitude of the clusters. Furthermore, a fine-grained understanding of the patient profiles is of utmost importance to employ them effectively when designing an application. In this paper, we suggest a graphical approach to visualize patient data, so a thorough understanding of the patient profile can be achieved. Figure 6 gives an example of some patients that we had a closer look at in Section 2.

When constructing the visualization, several decisions need to be made. First of all, it should be decided **which data** we want to visualize (i.e. the data represented in Fig. 6). Given the growing number of data collected by applications, it will be often unfeasible to visualize all data that we collect about patients. Therefore, it is important to choose well which data we want to visualize, e.g. do we want to gain insight into the relationship between outcome variables (e.g. weight) and application usage (e.g. number of times using the feature to view the weight evolution) or do we want to gain insight in only outcome or application usage data?

After deciding upon which type of data we want to visualize, we should determine the **level of detail** we want in our data (i.e. the vertical axis in Fig. 6). Depending on the application at hand, we can choose to group data on an application module level (e.g. the physical activity module, the medication adherence module, and the education module) or on a feature level (e.g. for the physical activity module, the step tracking feature or the sports reporting feature).

Next, it should be decided over **which time frame** we want to visualize the patient data (i.e. the horizontal axis in Fig. 6). This can range from all time (from the start until the end of the data stream) to only a day or even a few minutes. Of course, the possible time frames that can be selected depend on the richness of the collected data. Dependent on the selected time frame, a different unit of time can be defined (e.g. visualize information on a daily vs. monthly basis).

Dependent on the chosen data, level of detail and time frame, it should be determined **which change in the chosen metric is relevant** to visualize (i.e. indicated as the color gradient in Fig. 6). For example, if we want to visualize patients daily usage of each application module of a health application, we use the number of times that the patient opens the application module on that day as the metric to visualize. We may want to visualize the number of uses exactly, or alternatively divide it in certain categories (e.g. low, medium, and high usage). Another option is to combine both approaches, e.g. visualize the number of uses exactly as long as it is below 20 uses on a daily basis.

Lastly, it is important to decide **how many patients** to include in your graphical analysis (i.e. the vertical axis in Fig. 6). For large data sets, it is not possible to visualize all patients at the same time. Therefore, a decision needs to be made upon the patients that you want to visualize. Depending on the size of the clusters, it might be more feasible to visualize all or only a subset of the patients. Also, visualizing key patients from different clusters is an option. In addition, it should be decided if you want to exclude patients or filter them based on a certain condition. For example, you may want to exclude patients that did not use the application extensively, or only have a look at patients with a good medication adherence. These decisions should be made based on the factors that differentiate the clusters from one another and/or the research question at hand.

By analyzing the resulting graphical representation, it is possible that researchers may want to further separate clusters based on certain factors or characteristics useful to the domain of interest. We believe this is valuable, since these parameters might be relevant to personalize the health application in a later stage.

In the graphical representation for HeartHab (Fig. 6), we visualize the usage data on a module level for the entire study duration (i.e. the duration of the study for the patient that participated the longest in the trial) for the 5 patients that were discussed in Section 2. Each day in the study is represented by one block. The darker the color of the block, the more the module was used by the patient on that day. Blocks in which the patient did not have access to the application are depicted in grey.

3.4 Describing patient profiles using personas

The graphical representation created in the previous step can provide researchers insight into the similarities and differences between different patient clusters. These might help them in constructing personas that are relevant and achieve their purpose of supporting the development of an improved health application. Salminen et al. [15] analyzed 31 templates for data-driven personas. Their analysis indicated that the templates varied greatly in their information richness. They could not find one general template for data-driven personas. Moreover, it is challenging to create such a template, due to the variety of outputs of different methods and divergent information needs of persona users. Therefore, we do not impose a certain template for the development of a persona, but rather suggest to choose a template that fits with the goal of using the persona as a means to design a personalized health application. When creating a new health application or improving an already existing application, the created personas can be used during the design and/or evaluation stage.

Based on the graphical representation for HeartHab (in Fig. 6), we created personas. In Section 2, we already gave some examples of such personas. PAT030 is a person that is not tech-savvy and as a result did not use the application extensively. For this patient profile, the app should be easy to use and only include features that the patient is interested in.

3.5 Personalizing a health application according to patient profiles

The graphical representation and personas that were developed in the previous steps can be used as input for the design and development of a personalized health application. Especially the inspection of the graphical representation can yield interesting factors to (further) personalize the assistive and persuasive features of a health application or to create a new personalized application from scratch. The four key motivations to personalize from Section 2 can be used to analyze the chart and detect similarities and differences across individual patients, patient groups, or patient clusters, highlighting directions for personalizing the health application.

We conclude this section by presenting some examples of personalization features that could be proposed based on the graphical representation (Section 3.3) and the personas (Section 3.4) for a health application supporting cardiac patients in pursuing a healthy lifestyle:

- The graphical representation indicates that there is a group of patients who always successfully achieve their exercise goals, but all of a sudden they do not succeed anymore in their goals. This could be caused by the development of a co-morbidity or injury (e.g. back problems). In this situation, the application should not encourage patients to increase physical activity, because this is not feasible and would frustrate them. It would be better to divert the patient's attention to another behavioural goal (e.g. eating more healthy).
- The graphical representation indicates that people consistently report sports activities on the same days (e.g. always play tennis on Monday and go biking on Saturday). If they skip one of those activities (e.g. the patient did not report tennis on Monday), they do not achieve their weekly exercise goal. The application could be updated as such that if the patient does not report sports on one of his/her personal sports days, it prompts the patient to plan a sports session for another day.
- The graphical representation shows that patients always open the medication adherence module in the morning. This might be the ideal time to give a suggestion.

4 DISCUSSION

The use of personalized mobile interventions for lifestyle behaviour is still in its infancy [19]. Therefore, we investigated in this paper how to personalize applications targeting health behaviour change. In this paper, we made the following contributions:

(1) Define four key motivations for the personalization of health applications;
(2) Propose an approach to personalize health applications based on prior collected qualitative and quantitative study data, including an extra step using a graphical representation to analyze data collected by health applications to understand patient profiles and detect opportunities for personalization.

One of the major limitations of our work is the limited sample size and the extent of the log data that we used for the HeartHab examples described in this paper. Nevertheless, we believe that the examples and visualizations presented are useful for larger populations and provide interesting research directions on the synergy of human-computer interaction and the fields of big data analytics, machine learning, and artificial intelligence (e.g. clustering).

Recently, there is a trend towards interactive persona systems, in which persona users can interact with personas e.g. by choosing the data from which the personas are generated [14]. Although our presented approach to personalize health applications relies on data collected in prior studies, it is possible to use our approach in an interactive and dynamic fashion. Data can be analyzed at the moment of collection. This can be done automatically by using machine learning and artificial intelligence but also manually by using approaches such as our graphical representation and interactive personas. We can even go one step further and go to a fully automatic approach that dynamically updates the health application in real-time. For this to be possible, the application should be defined in a configurable or rule-based fashion, such that complete modules or features can be activated, deactivated or altered.

5 CONCLUSION

Personalization can be employed in health applications to motivate patients for behaviour change. However, till data, the use of personalization in health applications targeting behaviour change is limited. Based on data collected in our former studies, we identified four key motivations to personalize applications supporting health behaviour change. We presented a mixed-methods approach for the design of a personalized health application that uses qualitative and quantitative data to identify patient profiles and highlight directions for personalization. The main difference between our proposed approach and other existing approaches (e.g. data-driven personas) is the use of a graphical representation to understand patient profiles and identify ways to personalize health applications. We hope that our research can be a first step in making big data sets collected from studies easier to understand and usable for building successful health applications.

ACKNOWLEDGMENTS

This research was funded by the Special Research Fund (BOF) of Hasselt University (BOF18DOC26) and the EU funded project H2020 IA CoroPrevention (848056).

REFERENCES

[1] Bushra Alsaadi and Dimah Alahmadi. 2021. The Use of Persona Towards Human-Centered Design in Health Field: Review of Types and Technologies. In *2021 International Conference on e-Health and Bioengineering (EHB)*. 1–4.

[2] Marco Ambrosetti, Ana Abreu, Ugo Corrà, Constantinos H Davos, Dominique Hansen, Ines Frederix, Marie C Iliou, Roberto FE Pedretti, Jean-Paul Schmid, Carlo Vigorito, et al. 2021. Secondary prevention through comprehensive cardiovascular rehabilitation: From knowledge to implementation. 2020 update. A position paper from the Secondary Prevention and Rehabilitation Section of the European Association of Preventive Cardiology. *European journal of preventive cardiology* 28, 5 (2021), 460–495.

[3] John M Carroll. 2003. *Making use: scenario-based design of human-computer interactions*. MIT press.

[4] Haiyan Fan and Marshall Scott Poole. 2006. What is personalization? Perspectives on the design and implementation of personalization in information systems. *Journal of Organizational Computing and Electronic Commerce* 16, 3-4 (2006), 179–202.

[5] Carlos A. Gomez-Uribe and Neil Hunt. 2016. The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Trans. Manage. Inf. Syst.* 6, 4, Article 13 (Dec. 2016), 19 pages.

[6] Mieke Haesen. 2011. User-Centered Process Framework and Techniques to Support the Realization of Interactive Systems by Multi-Disciplinary Teams.

[7] Richard J. Holden, Anand Kulanthaivel, Saptarshi Purkayastha, Kathryn M. Goggins, and Sunil Kripalani. 2017. Know thy eHealth user: Development of biopsychosocial personas from a study of older adults with heart failure. *International Journal of Medical Informatics* 108 (2017), 158–167.

[8] Paulus Kirchhof, ESC CRT R&D, and European Affairs Work Shop on Personalized Medicine et al. 2014. The continuum of personalized cardiovascular medicine: a position paper of the European Society of Cardiology. *European Heart Journal* 35, 46 (08 2014), 3250–3257.

[9] Predrag Klasnja, Shawna Smith, Nicholas J Seewald, Andy Lee, Kelly Hall, Brook Luers, Eric B Hekler, and Susan A Murphy. 2019. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. *Annals of Behavioral Medicine* 53, 6 (2019), 573–582.

[10] James MacQueen et al. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Vol. 1. Oakland, CA, USA, 281–297.

[11] Steve Mulder and Ziv Yaar. 2006. *The user is always right: A practical guide to creating and using personas for the web*. New Riders.

[12] John Pruitt and Tamara Adlin. 2006. *The Persona Lifecycle : Keeping People in Mind Throughout Product Design*. Morgan Kaufmann.

[13] John Pruitt and Jonathan Grudin. 2003. Personas: practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences*. 1–15.

[14] Joni Salminen, Kathleen Guan, Soon-gyo Jung, Shammur A Chowdhury, and Bernard J Jansen. 2020. A literature review of quantitative persona creation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[15] Joni Salminen, Kathleen Guan, Lene Nielsen, Soon-gyo Jung, and Bernard J Jansen. 2020. A Template for Data-Driven Personas: Analyzing 31 Quantitatively Oriented Persona Profiles. In *International Conference on Human-Computer Interaction*. Springer, 125–144.

[16] Supraja Sankaran, Cindel Bonneux, Paul Dendale, and Karin Coninx. 2018. Bridging patients' needs and caregivers' perspectives to tailor information provisioning during cardiac rehabilitation. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32*. 1–11.

[17] Supraja Sankaran, Paul Dendale, and Karin Coninx. 2019. Evaluating the impact of the HeartHab app on motivation, physical activity, quality of life, and risk factors of coronary artery disease patients: multidisciplinary crossover study. *JMIR mHealth and uHealth* 7, 4 (2019), e10874.

[18] Supraja Sankaran, Ines Frederix, Mieke Haesen, Paul Dendale, Kris Luyten, and Karin Coninx. 2016. A grounded approach for applying behavior change techniques in mobile cardiac tele-rehabilitation. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 1–8.

[19] Huong Ly Tong, Juan C. Quiroz, A. Baki Kocaballi, Sandrine Chan Moi Fat, Kim Phuong Dao, Holly Gehringer, Clara K. Chow, and Liliana Laranjo. 2021. Personalized mobile technologies for lifestyle behavior change: A systematic review, meta-analysis, and meta-regression. *Preventive Medicine* 148 (2021), 106532.

[20] Christos Troussas, Akrivi Krouska, and Cleo Sgouropoulou. 2020. Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education. *Computers & Education* 144 (2020), 103698.

[21] S. Vosbergen, J.M.R. Mulder-Wiggers, J.P. Lacroix, H.M.C. Kemps, R.A. Kraaijenhagen, M.W.M. Jaspers, and N. Peek. 2015. Using personas to tailor educational messages to the preferences of coronary heart disease patients. *Journal of Biomedical Informatics* 53 (2015), 100–112.

[22] Bernhard Wöckl, Ulcay Yildizoglu, Isabella Buber, Belinda Aparicio Diaz, Ernst Kruijff, and Manfred Tscheligi. 2012. Basic senior personas: a representative design tool covering the spectrum of European older adults. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*. 25–32.

[23] World Health Organization. 2021. Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) Accessed on 09/12/2021.

[24] Lucy Yardley, Bonnie Spring, Leanne Morrison, David Crane, Kristina Curtis, Gina Merchant, Felix Naughton, and Ann Blandford. 2016. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. *American Journal of Preventive Medicine* 51 (11 2016), 833–842.

[25] Lucia Yu, Ethan Benjamin, Congzhe Su, Yinlin Fu, Jon Eskreis-Winkler, Xiaoting Zhao, and Diane Hu. 2021. Personalization in E-commerce Product Search by User-Centric Ranking. (2021).

[26] Xiang Zhang, Hans-Frederick Brown, and Anil Shankar. 2016. Data-driven personas: Constructing archetypal users with clickstreams and user telemetry. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 5350–5359.

[27] Haining Zhu, Hongjian Wang, and John M Carroll. 2019. Creating Persona Skeletons from Imbalanced Datasets-A Case Study using US Older Adults' Health Data. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 61–70.