Commercial off-the-shelf consumer health informatics interventions: recommendations for their design, evaluation and redesign

Jenna L Marquard,1 Teresa Zayas-Cabán2

ABSTRACT

Objective The goal of this paper is to describe the successful application of a use case-based evaluation approach to guide the effective design, evaluation and redesign of inexpensive, commercial, off-the-shelf consumer health informatics (CHI) interventions. Design Researchers developed four CHI intervention use cases representing two distinct patient populations (patients with diabetes with high blood pressure, post-bariatric surgery patients), two commercial off-the-shelf CHI applications (Microsoft HealthVault, Google Health), and related devices (blood pressure monitor, pedometer, weight scale). Three patient proxies tested each intervention for 10 days. Measurements The patient proxies recorded their challenges while completing use case tasks, rating the severity of each challenge based on how much it hindered their use of the intervention. Two independent evaluators categorized the challenges by human factors domain (physical, cognitive, macroergonomic). Results The use case-based approach resulted in the identification of 122 challenges, with 12% physical, 50% cognitive and 38% macroergonomic. Thirty-nine challenges (32%) were at least moderately severe. Nine of 22 use case tasks (41%) accounted for 72% of the challenges. Limitations The study used two patient proxies and addressed two specific patient populations and low-cost, off-the-shelf CHI interventions, which may not perfectly generalize to a larger number of proxies, actual patient populations, or other CHI interventions. Conclusion CHI designers can employ the use case-based evaluation approach to assess the fit of a CHI intervention with patients’ health work, in the context of their daily activities and environment, which would be difficult or impossible to evaluate by laboratory-based studies.

As authorized by the Health Information Technology for Economic and Clinical Health Act provisions within the American Recovery and Reinvestment Act of 2009, the Centers for Medicare and Medicaid Services has specified rules regarding the initial criteria that healthcare professionals and organizations must meet to qualify for specific monetary incentives (75 Fed. Reg. 44313). In particular, a portion of the Medicare and Medicaid Electronic Health Record Incentive Program addresses ‘patient engagement’ by requiring organizations to provide patients with a human readable and computable copy of their records. Stage 2 of meaningful use may require the use of personal health records—fostering heightened expectations about patients’ abilities to glean information from their providers’ records, maintain their own records, and share their records with various healthcare providers. As these national initiatives addressing patient engagement continue to expand, policy makers and healthcare providers must acknowledge that the home and other locations of daily living (LDL) such as workplaces, parks, exercise facilities and grocery stores are important sites for health and healthcare activities.1–3

Consumer health informatics (CHI) interventions based on inexpensive, off-the-shelf technologies that can easily be implemented in a wide variety of LDL are needed to cost-effectively support high quality, technology-enabled healthcare. Fortunately, lay people now have access to a variety of inexpensive devices such as weight scales, heart rate monitors and blood pressure cuffs. Many of these devices have the capacity to store user information over time and connect to consumer-directed applications, from which they may be able to share information with family members, friends and their care providers.

CHI intervention developers face a challenging task in designing simple, effective CHI applications and devices that support a diverse range of lay people using the applications and devices in a wide range of LDL.4 In addition, CHI interventions composed of inexpensive, commercial off-the-shelf CHI applications and devices will likely have more use challenges than controlled, institutionally developed CHI interventions or electronic medical record vendor-provided products, as off-the-shelf interventions link together several independent technologies.5 Resolving these use challenges will require the coordination of varied stakeholders. While some use challenges must be addressed by the commercial off-the-shelf CHI application and device designers, others can be addressed by supplemental intervention artifacts and processes or by patient training.

However, CHI intervention developers—who are often young, technology-savvy and without significant physical limitations—are largely taking a technology-centered design approach, paying minimal attention to lay people’s skills, knowledge and training, motivations, or psychosocial characteristics.6–11 CHI intervention developers often focus on automating existing tasks and rely on physicians’ perceptions of lay people’s problems and needs.12 Lay people are burdened by this approach, as it adds to their existing work, increases their frustration, and leads to errors, all of which may cause them to misuse and/or disuse CHI interventions.9 10 12 This technology-centered
design approach can reduce or negate the potential benefits of the CHI interventions, can increase developers’ investments, and can potentially compromise patient safety. Unfortunately, conceptual and methodological frameworks that effectively inform CHI intervention design and implementation are in flux because the field is growing quickly and is composed of researchers and practitioners from a variety of disciplines (eg, health communication, mass communication, information science and medical informatics). To improve the likelihood of the adoption and use of CHI interventions, developers need to understand lay people’s health work and the extent to which the CHI interventions support that health work. Although some studies have examined the usability of CHI applications by laboratory-based studies and heuristic evaluation, few have evaluated how CHI applications can best support lay people’s health work in context. 

The goal of this paper is to apply a use case-based human factors evaluation approach developed by Zayas-Cabán and Marquard to evaluate patient and provider proxies’ use of inexpensive, commercial off-the-shelf CHI interventions aimed at supporting two specific lay populations for which self-care and disease management are important, and in which poor outcomes can be costly—individuals with diabetes with high blood pressure and post-bariatric surgery patients. In this way, the goal of this paper is to serve as a feasibility study addressing the value of the proposed use case-based human factors approach. While many CHI interventions might aid these populations (eg, portals provided by electronic health record vendors or institutionally developed portals such as PatientSite and Patient Gateway), the CHI interventions examined in this study link inexpensive devices (eg, blood pressure cuff, weight scale, pedometer) with free applications (eg, Microsoft HealthVault, Google Health) that aggregate information from the devices and other data sources, such as commercial pharmacies and patient portals. Lay people can share information from the applications with family members, friends and providers, and the applications can be connected directly to their provider’s electronic health records. In this study, we focus on lay people who are actively engaged with a healthcare provider (hereafter referred to as patients).

BACKGROUND

Zayas-Cabán and Marquard posit that to be effectively adopted and used, CHI interventions must fit lay people’s physical, cognitive and organizational (or macroergonomic) ergonomic needs and constraints, and they provide use cases describing how these factors constrain patients’ abilities to use CHI interventions. In the context of this study, in which physical ergonomics encompasses the physical characteristics of patients and their work environments, cognitive ergonomics focuses on aiding patients as they process information while completing health work alone or in teams. Macroergonomics addresses the broader types of health work patients engage in (eg, disease management), their workflows (eg, the flow of health information across space and time, interactions with caregivers across space and time), and their work systems (eg, the social and organizational conditions under which their health work is performed). In the context of a patient self-monitoring blood pressure, physical ergonomics considers factors related to the patient operating the device (eg, securing the cuff) within the context of a physical environment (eg, at the kitchen table). Cognitive ergonomics considers factors related to processing information from the device’s user interface (eg, interpreting systolic and diastolic blood pressure values). Macroergonomics considers the context within which the device is used (eg, the patient should rest for a few minutes before taking a reading). Design attributes may affect multiple human factors domains. For instance, small print on the monitor interface limits patients’ ability to read the monitor values and consequently limits their ability to interpret and take action based on the readings. Their paper describes how the framework can inform the design, evaluation and redesign of inexpensive, off-the-shelf CHI interventions.

As CHI interventions are often targeted at patients with chronic diseases—frequently the elderly—patients may possess a variety of diminished physical characteristics (eg, reduced vision, hearing, dexterity and motor function). A patient’s use of CHI interventions may also be impacted by his/her decision-making style, diminished cognitive function, language, literacy level, perceptions of his/her current health status and predictions about his/her future health status. The goal of this paper is to serve as a feasibility study addressing the value of the proposed use case-based human factors approach, this paper focuses on two patient populations (individuals with diabetes with high blood pressure and post-bariatric surgery patients), two sets of CHI applications (Microsoft HealthVault and Google Health) and related devices (blood pressure cuff, pedometer, weight scale).

METHODS

The research team developed patient use cases, with the aid of healthcare providers, for each patient population/CHI application and device dyad. We feel it is essential that we compare the results of our study using patient proxies with results from...
that simulated the vision of a patient with cataracts.23

loaded pictures of the device and system screenshots into a tool to simulate the 73-year-old patient’s provider interactions. We provided the proxies with the detailed tasks in which we can foresee being able to make this comparison. These use cases were therefore developed to inform full-scale CHI intervention implementations for projects in which the authors are involved. The use cases in tables 1 and 2 detail activities that patients will likely undertake as they use the CHI interventions, including devices, applications and interpersonal interactions.

Use case wording that differs between consumer health informatics interventions for a specific patient population is shown in bold.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Use cases provided to patient proxies with diabetes and high blood pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HealthVault</strong></td>
<td><strong>Google Health</strong></td>
</tr>
<tr>
<td>Use case</td>
<td>You are a 73-year-old type 2 diabetic with high blood pressure. You are currently taking glipizide 5 mg once a day before breakfast for your diabetes and lisinopril 10 mg once a day for your high blood pressure. You also have cataracts and have arthritis in your hands. Your primary care physician would like you to take your blood sugar and blood pressure readings daily so that a nurse can change your medication level if your blood sugar or blood pressure levels get too high or too low. You have been given a blood pressure cuff to take these readings at your home. You are told to take the readings (1 \times ) per day at the same time of day, and wait at least 15–30 min after exercising. Your doctor also tells you to connect the blood pressure cuff to your computer and upload the readings from the cuff into an online system called Microsoft HealthVault. When you upload your blood pressure readings, your doctor would also like you to type in your blood sugar readings and your medications using HealthVault, and share your record with him/her. Your nurse will contact you if (s)he wants you to change your medication type or dose.</td>
</tr>
</tbody>
</table>

Use case wording that differs between consumer health informatics interventions for a specific patient population is shown in bold.

actual patient users. For this reason, we chose to target use cases in which we can foresee being able to make this comparison. These use cases were therefore developed to inform full-scale CHI intervention implementations for projects in which the authors are involved. The use cases in tables 1 and 2 detail activities that patients will likely undertake as they use the CHI interventions, including devices, applications and interpersonal interactions.

Use case wording that differs between CHI interventions for a specific patient population is shown in bold. Unlike usability studies that focus solely on discrete tasks, these use cases describe how the intervention is to be used over time in the context of the patients’ environment and existing self-care work processes. Along with each use case, the patient proxies were given a set of tasks required to successfully complete the use case. Box 1 provides a consolidated list of these tasks.

Three students—patient and provider proxies—tested each CHI intervention for 10 days. Two students performed ‘patient tasks’, while the other student, a registered nurse, performed ‘provider tasks’ to provide the patient proxies with realistic provider interactions. We provided the proxies with the detailed use cases and sets of related tasks to help them achieve their goals. In the diabetes use case, the patient proxies also tested the system while wearing medium-weight gardening gloves to simulate the 73-year-old patient’s arthritis, and researchers loaded pictures of the device and system screenshots into a tool that simulated the vision of a patient with cataracts.25

Over the course of the 10-day trial, the patient proxies recorded the challenges they encountered while completing each task. The proxies carried out the tasks in their homes and other LDL such as public gathering spaces and hotel rooms. In this paper, we focus only on the patient proxies’ challenges as the provider proxy was not conducting the trial within the context of her clinical work.

At the end of the 10-day trial the proxies rated the relative severity of each challenge using categories developed by Sears and Hess—0 for no problem, 1 for low severity, 2 for medium severity, and 3 for high severity—to determine the possible impact of the challenges on the use of the system.86 These ratings are largely practical, in that they allow designers and implementers to address the most serious challenges first. For severity ratings in which two of three raters agreed (eg, ratings of 2, 2 and 3), we used the majority rating (eg, 2). For ratings that differed across all three reviewers (eg, 0, 1, 2), we used the rounded average of the three scores (eg, 1).

Two independent evaluators analyzed and categorized the challenges by human factors domain (physical, cognitive, or macroergonomic). They then jointly discussed and resolved any challenges on which they initially disagreed about the categorization. The evaluators also consolidated similar challenges to make the guidance most useful to designers, and separated challenges that spanned more than one human factors domain.

**RESULTS**

Using our evaluation approach, we were able to identify 122 unique patient challenges across use cases. Figure 1 details the number of physical, cognitive, or macroergonomic patient challenges identified for each use case, with the numbers in parentheses signifying how many of the challenges were severe. The inter-rater agreement was 78% for the HealthVault diabetes use case, 85% for the Google diabetes use case, 97% for the
HealthVault challenges were macroergonomic. For instance, we directed each patient to goals, but still allowed them flexibility in completing each task. The research team could have specified subtasks that would enable the patient proxies to complete this task, such as which icons the patient should click on to create an account. However, this fine level of detail takes almost all cognitive load and reasoning away from the patient, thus likely limiting insightful information the researchers would otherwise obtain about what challenges the patient encounters in completing the tasks. The researchers could have, similarly, specified only higher level activities, such as ‘Use (HealthVault/Google Health) to send your blood pressure readings to your nurse.’ This activity-level guidance would likely not provide the patient proxies with enough direction to elicit detailed information on what challenges the patient proxies encountered using various aspects of the system, and the elicited challenges may be so diverse that comparing and synthesizing challenges across the patient proxies would be difficult. Therefore, the approach used (ie, use cases with a set of tasks) is more structured than heuristic usability evaluation (ie, providing little guidance and minimal information regarding the patient proxies’ tasks), but more flexible than cognitive walkthrough approaches (ie, using step-by-step task descriptions).

The categorization of the challenges as physical, cognitive, or macroergonomic is useful, in that design changes to mediate these challenges should align with the type of problem patients are experiencing. For instance, one could extract all cognitive challenges from the data set to understand how the system is or is not supporting patients’ cognition and how the design could be changed or supplemented to enhance this support.

Physical challenges were most often linked to patients’ use of CHI intervention-related devices: both diabetes use cases involved a blood pressure cuff, both bariatric surgery use cases involved a weight scale, and the HealthVault bariatric surgery use case also involved a pedometer. The number of cognitive challenges was quite consistent across use cases, with most cognitive challenges associated with one-time activities such as learning to use the devices and applications. The number

### DISCUSSION

Our evaluation approach identified a significant number of challenges likely to interfere with patients’ abilities to use a particular CHI intervention. While this list is likely not exhaustive, the large number of identified challenges provides intervention designers with a starting place from which to begin redesigning the intervention. Intervention designers can thus resolve challenges before conducting further, higher-cost testing with patients. Actual patients will likely encounter other challenges in addition to those identified through the described methodology, but the additional challenges will become more salient and easier to identify once the initial set of challenges is mediated.

The use case tasks enabled the patient proxies to know their goals, but still allowed them flexibility in completing each task. For instance, we directed each patient to ‘Create a (HealthVault/Google Health) account.’ The research team could have specified subtasks that would enable the patient proxies to complete this task, such as which icons the patient should click on to create an account. However, this fine level of detail takes almost all cognitive load and reasoning away from the patient, thus likely limiting insightful information the researchers would otherwise obtain about what challenges the patient encounters in completing the tasks. The researchers could have, similarly, specified only higher level activities, such as ‘Use (HealthVault/Google Health) to send your blood pressure readings to your nurse.’ This activity-level guidance would likely not provide the patient proxies with enough direction to elicit detailed information on what challenges the patient proxies encountered using various aspects of the system, and the elicited challenges may be so diverse that comparing and synthesizing challenges across the patient proxies would be difficult. Therefore, the approach used (ie, use cases with a set of tasks) is more structured than heuristic usability evaluation (ie, providing little guidance and minimal information regarding the patient proxies’ tasks), but more flexible than cognitive walkthrough approaches (ie, using step-by-step task descriptions).

The categorization of the challenges as physical, cognitive, or macroergonomic is useful, in that design changes to mediate these challenges should align with the type of problem patients are experiencing. For instance, one could extract all cognitive challenges from the data set to understand how the system is or is not supporting patients’ cognition and how the design could be changed or supplemented to enhance this support.

Physical challenges were most often linked to patients’ use of CHI intervention-related devices: both diabetes use cases involved a blood pressure cuff, both bariatric surgery use cases involved a weight scale, and the HealthVault bariatric surgery use case also involved a pedometer. The number of cognitive challenges was quite consistent across use cases, with most cognitive challenges associated with one-time activities such as learning to use the devices and applications. The number

---

**Box 1 Tasks performed by the patient proxies during the use cases**

1. Learn to use the blood pressure monitor.
2. Take a blood pressure reading.
3. Learn to use the pedometer.
4. Learn to use the weight scale.
5. Create a (HealthVault/Google Health) account.
6. Log into your (HealthVault/Google Health) account.
7. Install HealthVault software to upload device readings.
8. Upload (BP/weight/step counter) data into HealthVault.
9. Enter blood pressure readings into Google Health.
10. Enter blood sugar readings into (HealthVault/Google Health).
11. Enter medication data into (HealthVault/Google Health).
12. Enter weight into Google Health.
13. Link MSN ‘My Wellness Center’ to HealthVault.
15. Link ‘LiveStrong’ to Google Health.
16. Enter food intake into (My Wellness Center/’LiveStrong’).
17. Enter activity data into (‘My Wellness Center’/‘LiveStrong’).
18. Export food and activity information to spreadsheet.
19. Email the spreadsheet to your provider.
20. Grant provider access to record.
21. Log out.
22. General.

**Figure 1** Number of challenges identified by the patient proxies for each use case, by human factors domain.
of macroergonomic challenges varied more across use cases, with most macroergonomic challenges associated with repeated activities such as uploading device data, and regularly manually entering blood pressure, blood sugar and food intake data.

While some patient challenges are specific to particular use cases, others span devices and applications independent of use cases (eg, HealthVault for both diabetes and bariatric surgery). Device and application designers are best suited to address these challenges. For instance, if a device vendor resolves all challenges associated with its blood pressure cuff, it is resolving a subset of challenges that will arise for all use cases involving that blood pressure cuff. In addition, intervention designers can use supplementary intervention artifacts and processes and training material to guide patients as the patients interact with the devices and applications.

Other challenges are inherent to the task, so the challenges span use cases regardless of which devices and applications patients use (eg, both HealthVault and Google Health for bariatric surgery). These challenges provide rich opportunities for innovators to bring new technologies to the CHI marketplace, which may drive down the costs of existing technologies. For instance, the development of innovative and standardized means to collect and log food intake can mediate the confusion and complexity patients experience when manually logging individual foods and estimating calorie counts for specific foods.

Limitations
Our study has several limitations. First, our study only included two patient proxies, which limits the extent to which we can generalize the findings from this study to the challenges that a larger number of proxies might encounter. However, the two proxies identified 122 unique patient challenges, suggesting that this method can be used to successfully identify challenges that patients will encounter when using CHI interventions. Second, while our goal was to identify a set of use challenges using patient proxies before evaluating the system with actual patients, we did not compare proxy findings with challenges faced by actual patients. We realize that the set of challenges identified by actual patients will likely differ from those identified in this study and plan to conduct this analysis in future work. Our study focused on two specific patient populations, which have unique characteristics and limitations. However, these two patient populations are quite diverse: one is focused on a prevalent set of chronic diseases (diabetes, high blood pressure), and the other is focused on recovery from an acute intervention (bariatric surgery). It is also unlikely that any challenge is actually particular to only one use case. For instance, we identified challenges associated with uploading blood pressure data through HealthVault’s connection center during the HealthVault diabetes use case, but this challenge would be present for any use case involving the blood pressure cuff and HealthVault. Finally, because the study focused specifically on CHI interventions based on low-cost devices and applications that are not jointly designed, the challenges associated with these commercial applications may not well represent the type of challenges associated with customized systems. However, the approaches we used could be generalized to evaluate and redesign any type of CHI intervention.

CONCLUSION
We identified a large number of unique challenges by the use case-based evaluation approach, allowing CHI intervention designers to systematically assess the physical, cognitive and macroergonomic ‘fit’ of a particular commercial off-the-shelf CHI intervention with patients’ health work. In particular, we identified challenges that would be difficult, if not impossible, to detect in a laboratory-based usability study (eg, errors that happen every fifth time a device is used and errors that occur when uploading device data without an internet connection). The extended testing period allowed patients to log challenges as they occurred in the context of their daily activities and environment. Some tasks may be challenging on a given day and not on another (eg, being busy during the time of day one usually checks his/her blood pressure). Other tasks may unexpectedly present challenges (eg, computer equipment failure). Finally, the extended trial allows the designers to capture patients’ learning curves.

While testing with patients is ideal, our use case-based evaluation allowed proxies to take on the roles of patients and providers, which allowed the research team to identify a wealth of potential use challenges. The research team can now mediate these challenges before conducting extensive patient testing.

Acknowledgments The authors thank Janice Genevro, PhD, David Meyers, MD, and Jon White, MD from the Agency for Healthcare Research and Quality and Amy Brady from the University of Massachusetts for their valuable advice on this document. The authors also thank Kavita Radhakrishnan, registered nurse, Dana Enevend and Noah Duffey for their assistance acting as patient and provider proxies and aiding in data analysis. Preliminary findings related to one of the four case studies are published in the Proceedings of the 2009 American Medical Informatics Association Annual Meeting.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


