University of Louisville
ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

5-2016

A comparative usability study of web-based personal health records.

Alexandra Doggett
University of Louisville

Follow this and additional works at: https://ir.library.louisville.edu/etd

Part of the Industrial Engineering Commons

Recommended Citation
Doggett, Alexandra, "A comparative usability study of web-based personal health records." (2016). *Electronic Theses and Dissertations.* Paper 2355.
https://doi.org/10.18297/etd/2355

This Master's Thesis is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

A COMPARATIVE USABILITY STUDY OF WEB-BASED PERSONAL HEALTH RECORDS

By

Alexandra Doggett
B.S., University of Louisville, 2015

A Thesis
Submitted to the faculty of the
University of Louisville
J.B. Speed School of Engineering
as Partial Fulfillment of the Requirements
for the Professional Degree

MASTER OF ENGINEERING

Department of Industrial Engineering

May 2016

A COMPARATIVE USABILITY STUDY OF WEB-BASED PERSONAL HEALTH RECORDS

Submitted by: ___________________________________
Alexandra Doggett

A Thesis Approved on ___________________________________
(Date)

by the Following Reading and Examination Committee:

____
Jason Saleem, Thesis Director

____
Suraj Alexander

____
Matt Bohm

ACKNOWLEDGEMENTS

I would like to thank Dr. Jason Saleem and Mr. Dustin Weiler for all of their guidance, assistance, and input throughout the project.

I would also like to thank Dr. Jason Saleem, Dr. Suraj Alexander, and Dr. Matt Bohm for serving on my thesis review committee.

ABSTRACT

Previous research on Personal Health Records (PHRs) has focused on applications that are "tethered" to a specific Electronic Health Record (EHR). However, this is a gap in research on the usability of unaffiliated, independent PHRs, as well as a gap in research on college-aged PHR users. Based on this gap in the literature, a single factor, within-subject experiment was conducted using 18 participants from the University of Louisville to determine if one PHR had superior usability and design. The testing included the completion of six tasks in three different PHRs.

Dependent variables included task time, mouse movement, mouse clicks, keystrokes, errors, and usability survey results. The Computer System Usability Questionnaire (CSUQ), a validated survey instrument, was used for the usability survey. The experiment concluded with a follow up interview. ANOVA testing was completed on the results to determine the significance in the difference of the means. Results showed that several, but not all, measures had statistically and significantly different means. These included three survey categories, errors, and keystrokes. Though the hypothesis that all measures would be significantly different was partially supported, it can be concluded that Microsoft HealthVault has better usability than its tested counterparts, HealthSpek and Health Companion. The findings of this study could be used both by people looking to start using a PHR in the future and in findings obstacles to adoption of PHRs.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS…… ii

ABSTRACT…… iii

TABLE OF CONTENTS…… iv

LIST OF FIGURES…… v

I. INTRODUCTION…… 1

II. LITERATURE REVIEW…… 5

A. Value of PHR Literature…… 5

B. Studies on Usability…… 7

C. Studies on Usability with Certain Groups…… 10

D. State of Current Literature…… 14

III. INSTRUMENTATION AND EQUIPMENT…… 16

A. Experimental Design…… 16

B. Selection of PHRs and Usability Testing Tasks…… 16

C. Morae Software…… 19

IV. PROCEDURES…… 22

A. Settings and Sample Size…… 22

B. Usability Testing Procedures…… 22

V. RESULTS…… 24

A. Analysis of Individual Tasks…… 24

B. Analysis of PHR Systems…… 26

VI. DISCUSSION OF RESULTS…… 29

A. Analysis of Individual Tasks…… 29

B. Analysis of PHR Systems…… 32

C. Analysis of Qualitative Interviews…… 34

VII. CONCLUSIONS…… 42

VIII. RECOMMENDATIONS…… 44

APPENDIX I…… 45

APPENDIX II…… 47

APPENDIX III…… 49

APPENDIX IV…… 52

APPENDIX V…… 55

APPENDIX VI…… 56

REFERENCES…… 59

VITA…… 62

LIST OF FIGURES

FIGURE 1 – Home Page of HealthVault…..…………….....……………..…………… 18

FIGURE 2 – Home Page of Health Companion…...………………...………………… 18

FIGURE 3 – Home Page of HealthSpek...…………….………….………....……………… 19

FIGURE 4 – HealthSpek Medication Screen……………..……….........……………… 28

FIGURE 5 – Overall Satisfaction Ratings………………………….....……………… 30

FIGURE 6 – Microsoft HealthVault Medication Screen……..………………………... 31

FIGURE 7 – Microsoft HealthVault Registration Screen……..…...……...…………… 36

FIGURE 8 – Location of Modules on Home Screen of Health Companion…………… 38

FIGURE 9 – Module Page of Health Companion…………………………...…...…… 38

FIGURE 10 – Entering Blood Pressure in HealthSpek…..……...……….....…….……… 40

FIGURE 11 – Location of Log Out for HealthSpek..………..……….....……...………… 40

I. INTRODUCTION

A Personal Health Record (PHR) serves as a means for an individual to track his or her own health information. According to the Connecting to Health Initiative, a Personal Health Record is “an electronic application through which individuals can access, manage, and share personal health information, and that of others for whom they are authorized, in a private, secure, and confidential environment,” (Personal Health Working Group, 2003). While an individual has always been capable of tracking his or her own health, a PHR offers the opportunity for a person to centrally store health information. Software packages and Web-based versions currently exist, and these PHRs can often be used in conjunction with physician and hospital Electronic Health Records (EHRs), which allows for additional support through record integration. This allows the user’s entered information to be easily accessible to his or her healthcare providers.

Since PHRs offer a variety of services, different users seek out PHRs for different functions, which creates different defined user groups for PHRs. Users can be grouped based on the functionality they are looking for in a PHR and also by their age. Through a PHR, users have the opportunity to monitor their own health, including allergies, medications, medical history, weight, and other functionalities. In addition, users have the opportunity to track other family members’ health. The age of a user can often indicate what functions they are most interested in, as well as how easy it is for them to learn how to use a PHR.

New government policies have affected the future of PHRs. With the adoption of the Meaningful Use program by the federal government, health programs can receive monetary incentives for adopting EHRs. By 2015, however, companies that have not adopted an EHR will be penalized. In addition, there are two stages to the program. In the first stage, incorporating a PHR with the EHR is optional. The second phase of the program requires health companies to offer their patients a PHR, which could increase the use of PHRs in the future (Kannry, Beuria, Wang, & Nissim, 2012).

While there is research on PHRs in the literature, the proposed study of comparative usability testing addresses two gaps in the literature. The first gap that this study addresses is the lack of studies on college-aged students using PHRs. Past studies have focused on select user groups using a PHR (Crouch, Rose, Johnson, & Janson, 2015; Czaja et al., 2014; Fuji, Abbott, & Galt, 2015; Gee, Paterniti, Ward, & Soederberg, 2015; Shimada et al., 2014), but none have looked at college-aged students. In fact, of all these studies, only one study included a participant under the age of 30. This is an important demographic to analyze for several reasons. First, most college students have more technological experience than older users and should therefore be indicative of what future users will be like. While many college students do not have serious health issues that need tracking, starting to use a PHR at this age could be beneficial for future health reports. This study also addressed a gap in the literature of comparative usability studies. As previously mentioned, there are studies completed looking at individual populations (Crouch et al., 2015; Czaja et al., 2014; Fuji et al., 2015; Gee et al., 2015; Shimada et al., 2014) and individual PHRs (Price, Bellwood, & Davies, 2015; Curtis, Cheng, Rose, & Tsai, 2011; Ozok, Wu, Garrido, Pronovost, & Gurses, 2014; Monkman, & Kushniruk, 2013; Segall et al., 2011; Sheehan & Lucero, 2015), but there have been no studies solely focusing on gaps between different PHRs and attempting to determine a top performing PHR. One study recently completed took the first step in looking at PHRs comparatively (Czaja et al., 2015) but this study investigated "tethered" PHRs, which is a class of PHRs that are connected to a particular EHR. The PHRs investigated as part of this thesis are independent PHRs, not integrated with any particular EHR vendor.

Comparative usability research is beneficial because each PHR is different, so individual usability tests do not generalize to all PHRs. In addition, most of the studies that have been done on PHRs are qualitative in nature, without do not have substantial statistical analyses, while the proposed study looked at both objective and subjective measures. The findings of this study could also provide future benefits. For example, the findings could be useful to those who are interested in selecting a PHR to track their own health information. It could also be beneficial for the PHRs reviewed and PHRs being developed by revealing common errors in design that inhibit the usability of PHRs.

The need for the proposed study has been established by the lack of current literature on the subject (see section II). Therefore, an experiment was designed to complete a comparative usability study between three web-based, untethered PHRs (i.e., PHRs not connected to a specific EHR). In order to better understand PHRs, the study compared three leading PHRs through systematic usability testing on a sample of college age students to determine if one PHR has superior design and usability. It was determined through objective usability-based dependent variables, including time to complete a task, mouse clicks, mouse movement, keystrokes, and the number of commission and omission. This helped to answer the question, "Is there a difference in the usability of three leading online, untethered, free Personal Health Records for college aged students?" In addition to the quantitative metrics mentioned, a survey was also included for subjective metrics and qualitative feedback. The group experimental study used 18 participants who will complete six tasks in each of three different PHRs. The results were then statistically analyzed using ANOVAs to determine if there is a statistically significant difference in errors, time to complete tasks, or other relevant variables. Based on variability of PHR use and design reported in previous research, the hypothesis is that there will be a significant difference for each of the dependent measures previously described and that Microsoft HealthVault will be the most usable product based on these measures. This hypothesis is based on Microsoft's extensive resources and company history. Significant differences in the dependent measures would indicate a difference in the usability across the three PHR systems.

II. LITERATURE REVIEW

Three areas of research were reviewed and analyzed for the research question of this proposed study: (1) studies on the values and benefits of using PHRs; (2) studies on PHR and usability; and, (3) studies on PHRs with different populations. All of these areas are related to the usability of PHRs and past research completed on them. The collective findings indicate the potential benefits of PHRs, the usability improvements that are needed, and the groups that have been studied, as well as verify the need for such a study as described above.

A. Value of PHR Literature

One area where research has been conducted on PHRs is in the area of benefits and value of PHRs. This area of research emphasizes the need for further research to be completed on PHRs, since there are many potential benefits to their use but also a lack in the adoption of them. A set of interviews completed by Spil and Klein indicated that many people consider the use of PHRs as risky because of their distrust of security within the systems (2015). A considerable sample size of 83 was used, but inconsistent numbers were used for different health systems that were looked at, and although they state that age and demographics were taken into account, no participant demographic information is provided (Spil & Klein, 2015). Analysis of the interviews was completed, but no quantitative metrics were captured. Overall, the results of the study express user concerns for using PHRs, but more details are needed if this study was to ever be replicated or critiqued.

Instead of determining why PHRs are not being widely used, Kaelber et al. (2008) assessed the monetary value of PHRs by developing a model looking at eight different applications of PHRs in order to determine the value of a PHR. The team created a four-step process to analyze PHRs, which included technology definition, data collection, evidence framework, taxonomy definition, evidence synthesis, and model development. A literature review was completed to determine which benefits should be analyzed. The study found that over a ten-year adoption period, PHRs could result in up to $13 billion net annual value for the United States through eliminated additional monitoring and sharing of medical information (Kaelber et al., 2008). However, this study does not provide information of what PHRs are included in the analysis, though the literature review implied that a census, or complete review, was completed. In addition, because the field is relatively new, many assumptions were made in the quantitative analysis, such as the value of sharing a medications list and the adoption rate of the PHR, which, as previously mentioned, does not appear to be steady (Spil & Klein, 2015). As more research is done on PHRs, it would be beneficial to compare actual savings of using PHRs compared to the predicted rates.

Still yet another approach to determine the value of a PHR looked at the benefits the PHRs could provide its users. One study took a random sample of 250 participants with PHRs and sent them e-mails to inform them about the herpes zoster vaccination. In addition, 250 participants without PHRs were sent letters in the mail with the same information. The goal of the study was to determine if the functionality of PHRs improved the vaccination rate. Although the study found a higher response rate in participants with a PHR (53 as opposed to 12 from the non-PHR group), there was not a significant difference in the intervention methods (p=0.99). However, the higher response rates from PHR users indicate that further testing could be completed to look at other ways that PHRs can help intervene in patient care (Otsuka et al., 2013). In addition, it also shows that this area could be another benefit of PHRs. This functionality of communicating reminders with the user could be especially helpful for college-aged students to remind them to set up appointments or receive flu shots. A study by Tenforde, Jain, and Hickner also calls for the investigation of PHRs in order to "maximize the clinical value of this tool" (2011). They completed an analysis of current literature to determine if PHRs would be beneficial for chronic illnesses; however, their results were inconclusive based on the lack of literature in this area (Tenforde et al., 2011). No quantitative analysis was done on the results. All of these studies show potential values of PHRs, though literature on the actual value of different functionalities within PHRs still needs to be analyzed and written.

B. Studies on Usability

While there is little literature on the current values of PHRs, there are a variety of studies on the usability of individual PHR systems. This type of literature, in general, tests a user group's ease of use with the system by having them complete tasks with the system and then analyzing different aspects of the system to determine how to make it more usable. Within usability testing, there a variety of both qualitative and quantitative methods that can help determine usability issues, which do not always reach the same conclusions. For example, usability testing completed by Segall et al. showed qualitatively that the participants struggled in completing the tasks, but when given a survey, they rated the system high on ease of use on a scale of 1 to 5 (2011). The study used 20 volunteers and had the participants complete nine tasks in HealthView, which is linked to Duke University's Health System. "Think-aloud" method was used, where users verbalize their thought process as they complete tasks. Qualitative data is used with this method, as think-aloud technique can interfere with performance metrics, such as the time it takes to complete tasks. In addition, a survey was given at the end of the study. While over 85% of users agreed with statements like "I can navigate the system easily," most users had some difficulty completing tasks (Segall et al., 2011). In addition, an average score of 3.9 out of 5 was given on characteristics for the system, such as learnability and consistency, which indicates good usability. This gap indicates the importance of seeking out various parameters and tests when doing usability testing (Segall et al., 2011). In addition, it shows the importance of using objective dependent variables that are not judged by the user, as the user does not always perceive their own struggles. The proposed study uses both qualitative measures and quantitative methods that are not subjectively scored by the user, helping to ensure the validity of the results.

This technique of using multiple methods was also used by Monkman and Kushniruk, who used heuristic evaluation coupled with usability task completion (2013). Heuristic evaluation uses a group of experts along with a set of usability heuristics to analyze the usability of a system. This method intends to predict the issues that main user groups will endure with the current system. The task completion used in this study is similar to that of the proposed study but used the "think-aloud" method to qualitatively interpret the usability issues. Neither method was successful in determining all of the usability issues for the PHR, but overall 15 problems were identified, again showing the importance of using multiple methods. This study involved four participants and could use a larger sample size coupled with quantitative analysis to further verify the results (Monkman & Kushniruk, 2013).

Another approach for usability testing besides testing the end product is to evaluate the system during development, called formative usability testing. One study, much like the study of Monkman & Kushniruk (2013), used a group experimental design with usability testing, "think-aloud," and convenience sampling to determine usability for a new health record in Alberta, Canada. This type of usability is helpful in preventing and fixing usability issues before the system is available to the public. The study found several issues that were fixed before introduction to the public. In addition, the article states that there is a lack of evaluation overall for PHRs and even suggests follow up studies for further research, such as comparative evaluation, which is the intent of the proposed experiment (Price et al., 2015).

While there are several standard usability testing techniques that are widely accepted and commonly used, such as heuristic evaluation and "think-aloud" technique, there has also been a push to use more original techniques. Sheehan and Lucero (2015) used an innovative approach by evaluating not only the system, but also the users, tasks, and functions. It targeted a specific type of PHR based on fall prevention for an elderly community. Snowball sampling was used to recruit participants over the age of 55 to participate. There were four different modules used to assess the usability of the site and included aspects of functional analysis (Sheehan & Lucero, 2015). This method also employed a mix of qualitative and quantitative metrics in its evaluation, similar to the study by Segall et al. (2011). Overall user satisfaction was high, with an average satisfaction rating of 1.58 on a 7-point scale with one being strongly agree. The task with the highest error rate occurred with self-monitoring aspects of the site, which is addressed as an area to be improved (Sheehan & Lucero, 2015). Self-monitoring areas are a key aspect of a PHR, so this is an issue that should be addressed in future updates to the system. Self-monitoring functions, such as daily weight and blood pressure information, was also a focus in the proposed study.

Solely descriptive studies are another approach that has been used in studies analyzing PHRs. Ozok et al. used a combination of observations, "think-aloud," surveys, focus groups, and interviews to determine usability issues with MySafe-T.net (2014). Descriptive statistics determined means and standard deviations for the survey results, which used the inverse of the 7-point used by Sheehan & Lucero (2015). This study used 22 participants, but only one participant under 30, again showing the lack of studies using college-aged students, and used convenience sampling by asking patients from a caregiver's office. The purpose of this study was to determine design flaws in the current system that hinder usability. Sixteen design issues were determined, all of which were determined via qualitative analysis. Three of the flaws were also supported by quantitative analysis of the survey results. However, this study also has some limitations. Since it only involved one PHR, the results are not generalizable to all PHRs (Ozok et al., 2014).

C. **Studies on Usability with Certain Groups**

Perhaps the most relevant and recent studies that have been done on PHRs are those that test PHR usability on specific samples with the goal of representing a specific target population. Unlike many of the previously mentioned literature, these studies are targeting a specific population and analyzing their PHR needs instead of overall usability. This is directly related to the proposed study, which examined PHR usability for college-aged individuals. Though college-aged students have yet to be studied for PHR usability, veterans (Crouch et al., 2015; Shimada et al., 2014), chronically ill (Gee et al., 2015), low-income (Czaja et al., 2014), and diabetic (Fuji et al., 2015) adults have been analyzed.

One study of interest looked at chronically ill patients using a PHR. This differs from previous studies and the proposed studies in that it sought participants who had already been using the system for at least two years. The goal of recruiting these users was to learn how everyday users used the PHR and determine areas of improvement, which could potentially serve as a follow up study for the proposed study. The participants were recruited based on a list of eligible patients from a doctor's office. All eligible patients were recruited but around half responded and remained eligible for the study. The methods included a semi-structured interview individually with each participant. The interviews revealed four major themes for the PHR, which included patient engagement and health self-management, access to and control over personal health data, promotion of productive communication, and opportunities for training and education. One thing most participants noted was that they experience difficulty navigating the system's design/layout. This study provided verification of the main uses of this PHR, but still had limitations and improvements for future studies. Like many of the other studies, the sampling was not random and the results are not generalizable. In addition, a narrow age range (50-65) was studied, so additional ages could be considered in future studies (Gee et al., 2015).

Another study combined experienced users and new users' input by giving participants an initial training and then interviewing them 3-6 months later. This study specifically used 59 patients with type 2 diabetes from two clinics selected based on medical condition. It analyzed Microsoft HealthVault, one of the PHRs of interest for the proposed study. After the required period, interviews were completed, and nine themes emerged from the interviews. Of these, three were positive and included increased awareness and behavioral changes, and six were negative, including "I would have used it if I were sicker," and privacy and security concerns (Fuji et al., 2015). These negative themes can be the subject of future studies and modifications for the system and also reiterate comments from the study completed by Spil and Klein (2015) when analyzing issues in adoption of PHRs. In addition, these themes can be compared to the findings of the proposed study.

Veterans are another group that has been individually studied using PHRs. The Veterans Health Administration has their own PHR called My Health*e*Vet. My Health*e*Vet is considered a tethered PHR because it is linked to the VA's EHR. One study served as a way to understand the reach of the PHR depending on pre-existing conditions. This relates to the proposed study because one of the intended benefits of the study is to better promote usage of PHRs within the college community. This study completed a cross-sectional analysis, which included all veterans who used the VA over a two-year period, which included over six million veterans (Shimada et al., 2014). This is significantly larger than any other PHR literature. The study made adjustments for sociodemographic factors, and the results that veterans with HIV, hyperlipidemia, and spinal cord injury were most likely to use the PHR were statistically significant ($P < .001$). Though these results cannot be generalizable to the overall population, it could have repercussions for future studies on other populations (Shimada et al., 2014). For example, a study done by Crouch et al.

indicated that the use of My Health*e*Vet in veterans with HIV is associated with better satisfaction when setting appointments and getting care and information (2015).

Another study on veterans was completed by Haggstrom et al. (2011). This study took a similar approach to the proposed study in that it used similar tasks and task time as a measure of efficiency, as well as interviews. The study used registration, login, prescription refill, tracking health information, and searching for health information as the tasks that participants completed. These tasks had predetermined time limits and target times. Qualitative data was also collected in the forms of observations and debrief interviews. Many participants struggled to complete the tasks, with three of the activities having less than 30% completion. The qualitative data proved helpful in determining future changes that could improve the usability of the system, such as making information more accessible (Haggstrom et al., 2011). The proposed study used a similar layout but used a survey between PHRs with an interview at the end, whereas this study used an interview between each task.

Perhaps the most relevant and similar study to the one proposed is a study completed by Czaja et al. (2015). The study used task analysis and literacy load analysis for three different PHRs, along with usability testing for 54 adults of low socioeconomic status to determine the usability of these systems. The literacy load analysis indicated a high level of technical vocabulary, while the task analysis laid out the steps required to complete the tasks asked of the participants. After these steps were completed by the researcher, usability testing, which included a background questionnaire, a health literacy tool, system-rating questions, and a PHR rating scale was conducted. Significant differences were observed for finding medical information and interpreting lab results (P

$\leq .001$), with System A being the least difficult to use. However, overall the findings indicate that PHRs could be difficult to use, especially for people of lower literacy and economic levels. The authors call for more studies with larger, more diverse user groups. In addition, convenience sampling was used in this study as well. Usability testing for all three systems was also done in one sitting, which could have affected the ease of use (Czaja et al., 2015). While Czaja et al. used a counterbalanced design, the proposed study randomized the order and use order as a block variable when analyzing using ANOVA.

D. <u>State of Current Literature</u>

While PHRs are an exciting new technology with many applications, there are still issues in the adoption and usability of these systems. There have been three main areas of research about PHRs that relate to the proposed study: the value of using PHRs, the usability of a single PHR, and the usability of PHRs for certain groups. Overall, these studies lack quantitative and statistically significant support, as most of them act as descriptive studies to determine the errors in usability in different systems. In addition, only two pieces of literature describe comparative studies with two or more PHRs. None of the studies have a sample that includes college-aged students, let alone one that focuses on them, and none of them use the proposed dependent variables that were used in the proposed study. In addition, this study focused on independent PHRs, as opposed to most studies, which test PHRs “tethered” to a specific EHR. For these reasons, the proposed study will add to the literature by providing unique results that will answer questions previously unanswered in the research community. While these areas are important for future research, these are not the only areas that could benefit from further studying. Most of the studies, excluding a cross sectional analysis of My Health*e*Vet (Shimada et al., 2014), have relatively small sample sizes and call for validation studies using larger groups. In addition, most of the studies used convenience sampling, which limits the generalizability of the results. Validation studies that incorporate random sampling would improve the internal validity of these experiments.

Usability testing that has been analyzed in the literature has used a variety of techniques, which serves as a strength in many cases. As shown by Segall et al. (2011), the use of one method, especially qualitative methods, are not always indicative of valid results. Limiting the study to only one measure disallows for further validation of the results. While internal validity due to subject characteristics is a weakness of the current PHR literature, most authors recognize this as a limitation in the articles. Since most of the articles evaluate individual PHR systems and none use the same PHR, the comparison of results cannot be applied since the systems vary. Another benefit of the proposed study is that it allows for comparison of the PHRs used in the study. With the lack of comparative PHR literature, this added significant findings to the research community that will help others better understand the uses and usability issues for PHRs within a sample of a new target population.

III. INSTRUMENTATION AND EQUIPMENT

A. Experimental Design

The design chosen was a single factor, within-subject design, where each participant completed the same tasks in each of the three PHRs. A within-subject is more practical than a between-subjects design in terms of resources required and time needed to conduct the study. In addition, the within-subject design allowed for comparative feedback from each participant in a post-experiment interview. The single factor (independent variable) was 'PHR-type', with three levels (each of the three tested PHRs). In order to control for a learning effect between PHRs potentially affecting task time, the order that participants used the PHRs was randomized across participants. Since there are six different order combinations for the PHRs, each order combination was randomized but used three times for the total of 18 participants.

B. Selection of PHRs and Usability Testing Tasks

In order to choose the three PHRs to be used in the study, initial research was completed using https://www.myphr.com/resources/choose.aspx to assess current web-based, untethered, free PHRs. Common tasks done in PHRs were chosen for the study and included the following tasks: registration, entering information about medications, allergies, blood pressure, weight, and family history. Tasks were chosen to be representative of information that college-aged students would be interested in entering into their own PHR. Concerns arose about the registration task because if the participant was not able to successfully register for the site, the participant would not be able to complete the rest of the tasks. Therefore, it was decided that adding a member to an already existing PHR would serve as an initial task, so that even if the participant could not register, they would still be able to complete other tasks in the system. Because of the nature of the adding member and registration tasks, these tasks remained first and last tasks respectively for each PHR and participant. The other four tasks were randomized both for the PHR and the participant.

Once the tasks were chosen, PHRs were selected based on the previously mentioned criteria and the additional requirement that all of the tasks be present in the PHR. After analyzing the current systems, the PHRs chosen were Microsoft HealthVault, Health Companion, and HealthSpek. The study used version 3.5.14.0.2 of Health Companion. Current version numbers of HealthSpek and HealthVault could not be found on either website. The websites were accessed in the fall (September through November) of 2015. In addition, the three systems had different layouts, which was also preferable for usability testing. The home screens of the three systems are shown below (Figures 1,2, and 3). This study received approval from the University of Louisville's Internal Review Board (IRB). See Appendix I for the IRB outcome letter.

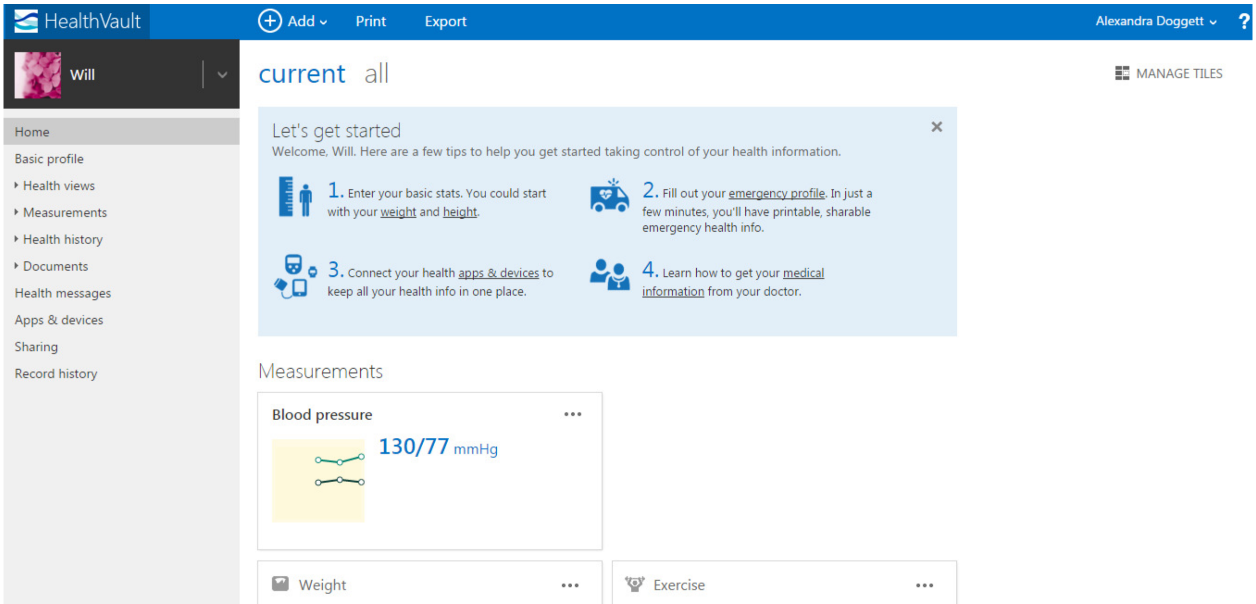

FIGURE 1 – Home Page of HealthVault

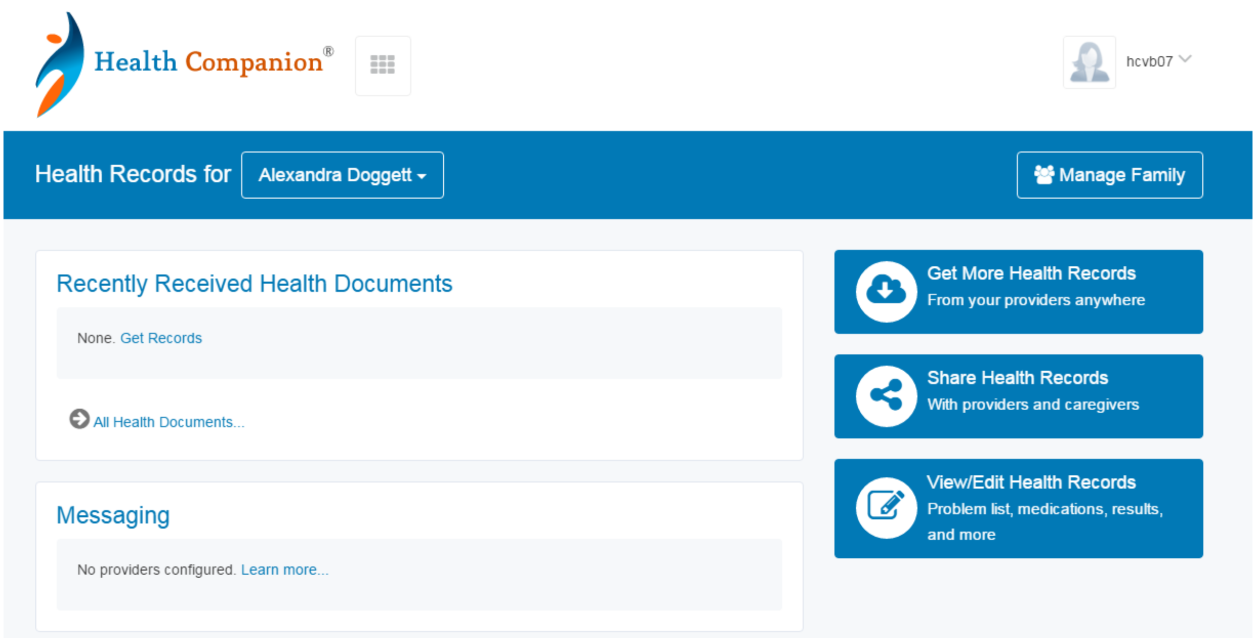

FIGURE 2 – Home Page of Health Companion

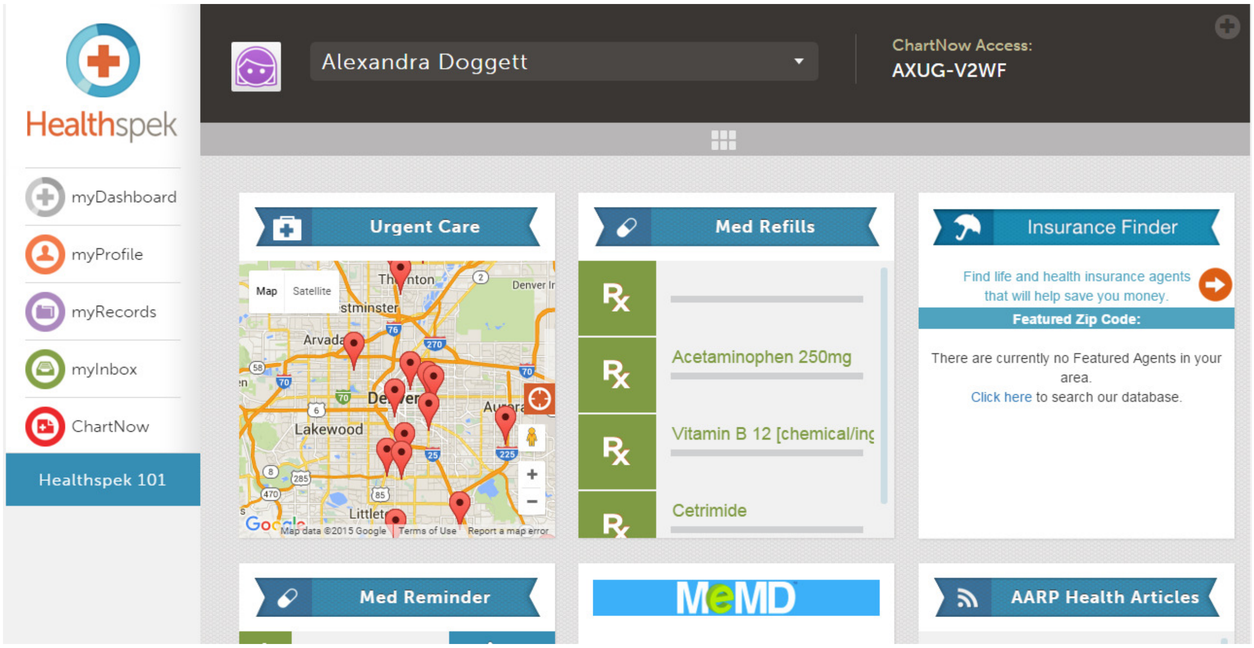

FIGURE 3 – Home Page of HealthSpek

C. Morae Software

Morae software (Version 3.1.1) by TechSmith was used to capture quantifiable data for the study. Morae is a software program that allows the user/experimenter to record and analyze both the computer screen and user as they complete tasks (Morae from TechSmith, n.d.). Morae was developed by TechSmith, a screen capture/recording software company founded in 1987. Morae, one of the software packages that the company now offers, was first produced in 2004 (Morae Version History, n.d.). It has since gone through 18 updated versions to improve its functionality. Morae has a variety of functions, such as software and web experience testing, focus groups, interviews, hardware testing, and paper prototype testing, which allows for a diverse user population, but mainly for testing for developed products for companies and usability testing (Learn How to Use Morae, n.d.).

The software captures a variety of variables, such as task time, keystrokes, mouse clicks, mouse movement, and comments from the experimenter. The comments can then be used to mark errors in the completion of tasks. The instrument is made up of three different items: Morae Recorder, Morae Observer, and Morae Manager. Morae Recorder is the interface used for the participant. It captures their screen and webcam images and displays the tasks to be completed. Morae Observer allows the experimenter to see the participant's screen and webcam and make comments during the actual testing. Morae Manager is used after the testing is finished and analyzes the variables previously specified.

Since the software program acts as the outer shell of testing and users develop their own tests/tasks to use with it, there is not significant literature on the validity or reliability of Morae. Its validity can be attested to via its popularity and the variety of customers who continue to use Morae, including Microsoft, Google, eBay, and Amazon.com (Case Studies and Customer Stories for Morae, n.d.). Morae has been used as an instrument in a significant amount of studies and published papers, including one that used it as a benchmark to compare to a new usability software (Sivaji, 2012), which is a form of the equivalent forms method for reliability. Although Morae has been used or referenced in over 150 scholarly articles (scholar.google.com), no published papers address the validity or reliability of the product, mainly due to the fact that the user is required to set up a test within the software.

The auto task logger function was used in Morae, meaning that instructions were given on the computer, and the participant chose when to start and end tasks. In addition, the Morae software system automatically displayed a survey at the end of participant's use of each PHR. The Computer System Usability Questionnaire (CSUQ) survey (Lewis,

1995) was chosen for the study. A study by Tullis and Stetson found that the CSUQ survey provided the same conclusions with a small sample size as those of a larger sample size 90% of the time (2004). In addition, CSUQ provides four different scales of measurement based on different groupings of questions, which include overall satisfaction, interface quality, information quality, and system usefulness. These additional measures offer additional opportunities to determine the likes and dislikes of the participants and narrow the usability issues.

IV. PROCEDURES

A. Settings and Sample Size

The experiment used 18 college-aged students from the University of Louisville for the study, all of whom had never used the PHRs in question. All students were between the ages of 18 and 24. In a study by Kushniruk & Patel, it was found that 8-12 participants could account for up to 80% of usability issues (2004). To ensure sufficient clean data for analysis, a higher number of participants was used. Convenience sampling was used to recruit the participants. Fifty percent of participants were male, and 50% were female. 83% were engineering students. All testing was done in the same lab and on the same computer to prevent internal validity threats due to location. The participant set up included a monitor, keyboard, mouse, and webcam, along with written instructions placed in front of the screen. All of the data was collected by the same data collector, who followed the script that can be found in Appendix II.

B. Usability Testing Procedure

Upon arrival at the lab, participants were asked to set their belongings down and were given a brief overview of the study. Participants were then asks to read and sign a consent form. See Appendix III for the full consent form. Once the form had been read and signed by the participant, the participant was led to the study room. The typed instructions along with the information to enter were set in front of the computer. These instructions, along with instructions on how to interact with the Morae software, were explained to the participant. Each participant was told to read over the instructions and information and to ask any questions before starting. These instructions can be found in Appendix IV. The data collector was in a room adjacent to the participant and watching the screen using Morae Observer to answer questions or direct the participant. Upon completion of the study, each participant was asked follow up interview questions. These questions can be found in Appendix V. Upon completion of these steps, participants were thanked for their time and dismissed.

V. RESULTS

A. Analysis of Individual Tasks

Individual tasks were analyzed using four different metrics: keystrokes, mouse movement, mouse clicks, and task time. These metrics were then compared across the three different PHRs. In addition, individual tasks had to meet several requirements in order to be considered complete. It was predetermined, such that the overall experimental session would not exceed 90 minutes, that no task should take over five minutes. Therefore, any participant that took over five minutes on a task was asked to move on to the next task. In addition, it was determined that any task that had more than five omission errors was also considered incomplete. These criteria, along with the participant actually doing the correct task, were used to determine if a task should be included in the descriptive statistics of individual tasks. In addition, ANOVAs were run on the data to determine the statistical significance of the findings. Table 1 below shows the completion rates for the 18 participants for each task.

TABLE I

NUMBER OF PARTICIPANTS THAT COMPLETED EACH TASK ON EACH PHR

Task	**Health Companion**	**HealthSpek**	**HealthVault**	**Total**
Add Member	16	14	16	46
Allergies	14	17	16	47
BP/Weight	15	16	7	38
Family History	16	17	18	51
Medications	15	3	18	36
Registration	14	16	12	42

Most tasks had at least 2/3 completion rate, with the exception of HealthSpek medications and HealthVault registration. In addition, HealthVault was the only PHR that had 100% completion in any task, which occurred in family history and medications.

When completing statistical analysis on the task data, ANOVAs were the first test used. An ANOVA is used to test if there is a significant difference between the two or more means. A significance level of α = 0.05 was used for all tests. If the data did not meet the normality assumptions necessary for an ANOVA, a Box-Cox transformation was applied to the data. If this still did not yield normal results, a Kruskal-Wallis nonparametric test was used on the data. Using these tests, it was determined that there was a significant difference in the means of the following tasks with the following metrics (Table 2):

TABLE II

SIGNIFICANCE OF DIFFERENCE OF MEANS FOR INDIVIDUAL TASK ANALYSIS

	Task Time	**Number of Clicks**	**Mouse Movement**	**Keystrokes**
Add Member	No Significant Difference	Significant Difference (Box-Cox Transformation, ANOVA)	No Significant Difference	No Significant Difference
Allergies	Significant Difference (ANOVA)	Significant Difference (ANOVA)	Significant Difference (Box-Cox Transformation, ANOVA)	Significant Difference (ANOVA)
Blood Pressure & Weight	Significant Difference (Box-Cox Transformation, ANOVA)	Significant Difference (Box-Cox Transformation, ANOVA)	Significant Different (Kruskal-Wallis)	No Significant Difference
Family History	No Significant Difference	Significant Difference (Kruskal-Wallis)	Significant Difference (Box-Cox Transformation, ANOVA)	Significant Difference (Kruskal-Wallis)
Medications	Significant Difference (ANOVA)	No Significant Difference	Significant Difference (Box-Cox Transformation, ANOVA)	Significant Difference (Box-Cox Transformation, ANOVA)
Registrations	Significant Difference (ANOVA)	Significant Difference (ANOVA)	Significant Difference (ANOVA)	No Significant Difference

As illustrated by the table, none of the metrics had significant results across all six tasks. There was, however, a significant difference across all metrics for the allergy task. A Tukey test was used to determine which means were significantly different for means that proved significantly different. In all of the metrics, HealthSpek was significantly different from Health Companion, and HealthSpek was significantly different from HealthVault in all but mouse movement. HealthSpek had higher average keystrokes and mouse movement and lower mouse clicks and task time for allergies.

B. Analysis of PHR Systems

After analyzing the individual tasks for each system, the PHR systems as a whole were analyzed. Metrics that were analyzed for the PHR systems included total task time, mouse movement, total keystrokes, total mouse clicks, total keystrokes, system usefulness, information quality, interface quality, overall satisfaction, total omission errors, and total commission errors. System usefulness, information quality, interface quality, and overall satisfaction are four subscales of the CSUQ survey and were determined based on the average of the relevant CSUQ questions. The significant results are shown in Table 3.

TABLE III

COMPARISON OF OVERALL RESPONSES FOR THREE DIFFERENT PHRs, n=18

Response	Factor(s)	HealthVault	Health Companion	HealthSpek	Significance
Task Time (s)	Order, Gender	975.1 (156.8)	1085.1 (243.4)	975.6 (273.5)	1 > 2, 3 Male > Female
Mouse Movement (pixels)	Order	137283.6 (39803.7)	173374.9 (59062.9)	159616.5 (48843.8)	1 > 2, 3
System Usefulness (1-7 scale, 7 high)	Order, System	5.78 (0.75)	4.97 (1.13)	4.59 (1.12)	3 > 1, HV > HS, HC
Overall Satisfaction (1-7 scale, 7 high)	Order, System	5.57 (0.75)	4.85 (1.10)	4.36 (1.14)	3 > 1, HV > HS
Information Quality (1-7 scale, 7 high)	System	5.35 (0.83)	4.66 (1.24)	4.00 (1.30)	HV > HS
Omission Errors	System	12.44 (4.59)	5.56 (5.73)	9.33 (4.55)	HV > HC
Comission Errors	System	0.17 (0.38)	1.56 (1.76)	0.94 (1.11)	HC > HV
Keystrokes	System	533.6 (107.5)	610.1 (109.1)	721.6 (210.9)	HS > HV

HealthVault was statistically significantly better on the survey metrics, which included overall satisfaction, system usefulness, and information quality. It also was scored the best on interface quality, though the results were not statistically significant (P = .098).

Another metric of interest that was analyzed in more detail was the overall satisfaction for each user and each PHR. This statistic was calculated by taking the average of all 19 questions on the CSUQ. These questions can be found in Appendix VI. This included a total of 54 points, one for each PHR for each participant. Figure 6 shows the overall satisfaction ratings of each PHR by participant.

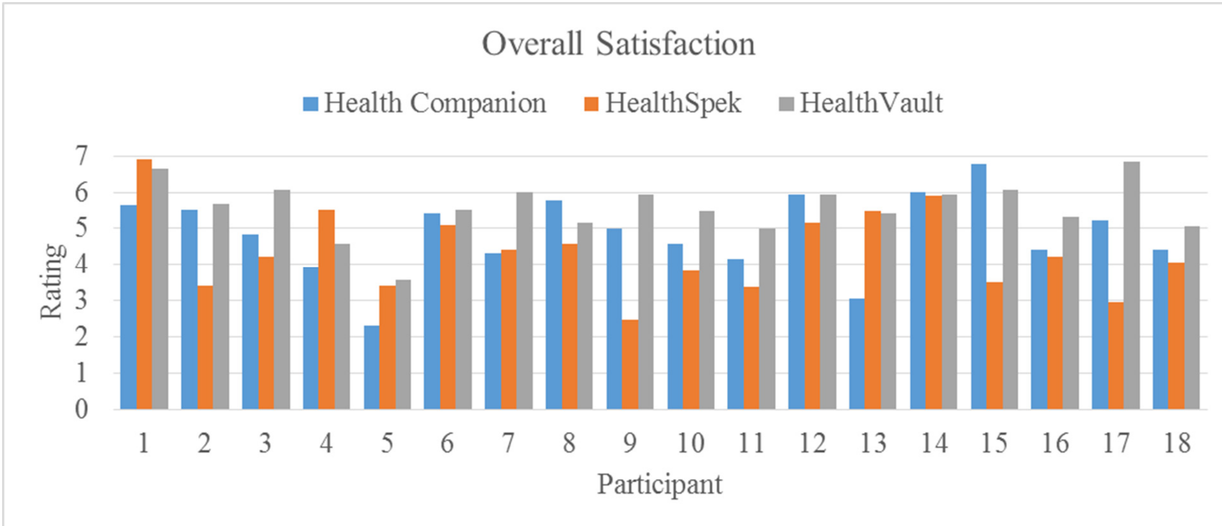

FIGURE 4 – Overall Satisfaction Ratings

VI. DISCUSSION OF RESULTS

A. Discussion of Individual Tasks

As supported by Table 1, most of the tasks in each system were completed by a majority of the participants. Two individual tasks stood out as outliers. The first was the HealthSpek medication task. The page to enter medication information was considerably different from the other two systems and contained more fields. Because of this, six of the participants were unable to complete the task within the time limit. The site also autocorrected to the first item typed in each field, causing more commission errors than other sites. This resulted in only three participants being able to complete the task within the determined criteria. The medication input screens for HealthSpek, compared to the medication input screen for Microsoft HealthVault, are shown in the images below in Figures 4 and 5.

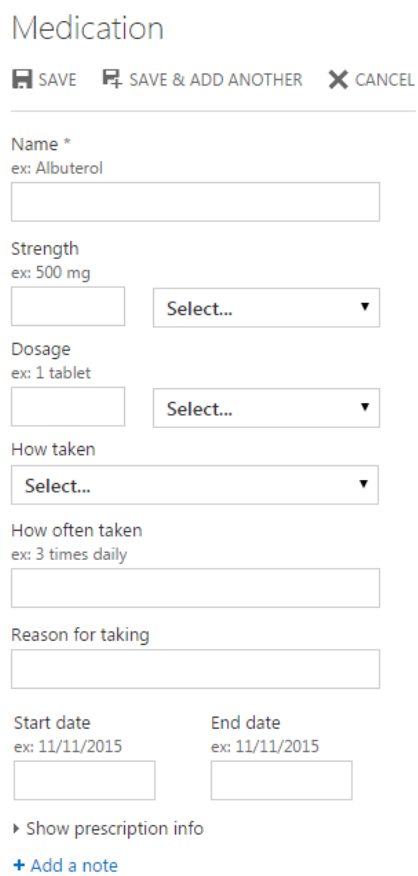

FIGURE 5 – Microsoft HealthVault Medication Screen

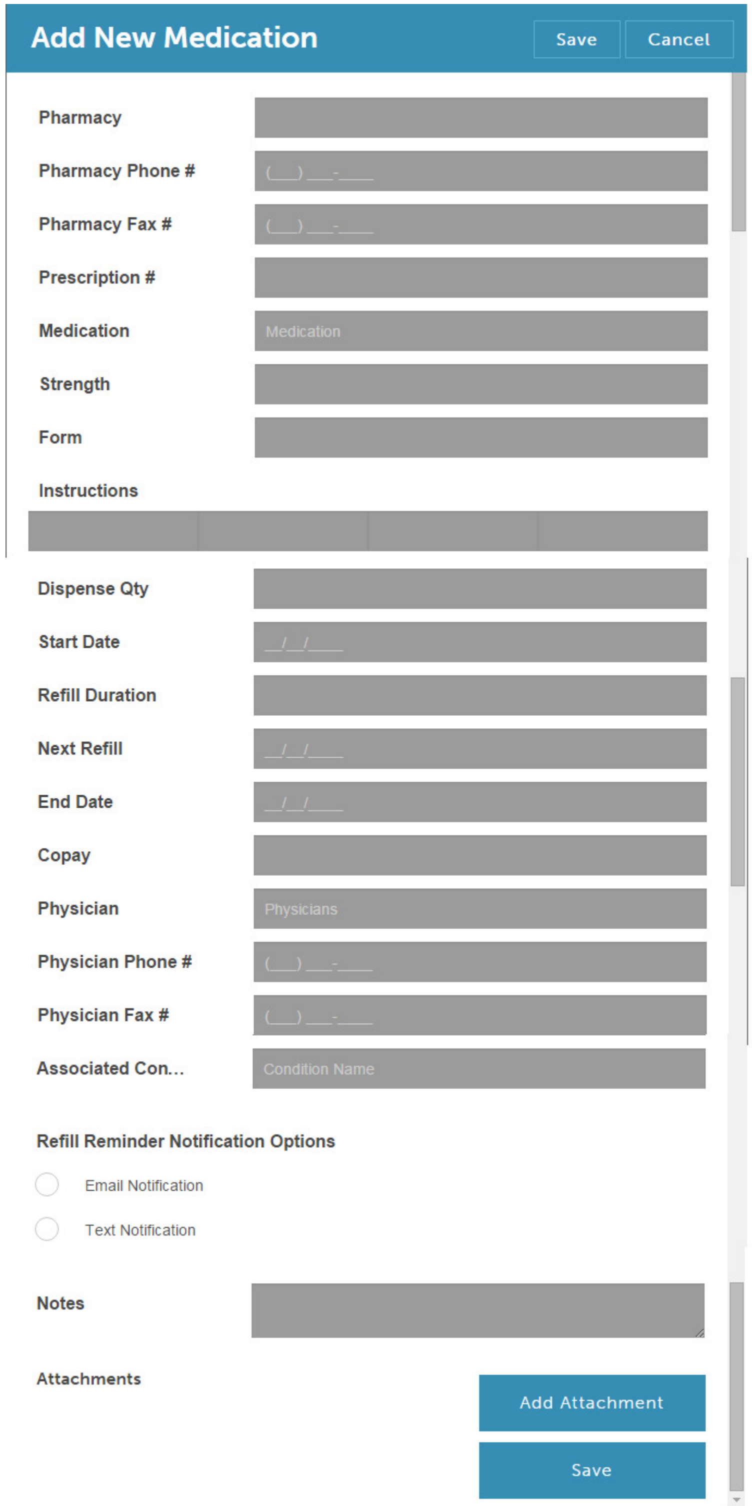

FIGURE 6 – HealthSpek Medication Screen

The other outlier was the HealthVault blood pressure and weight information task. The other two systems had these linked in one input page titled Vitals. However, HealthVault has these grouped separately. Because of this, many participants did not realize that they had to click somewhere else to enter the rest of the information and thus did not, resulting in only seven of the participants completing the task.

Table 2 findings indicated that HealthSpek had statistically significant higher average keystrokes and mouse movement and lower mouse clicks and task time for allergies. This is consistent with the participants' reflections in the post-experiment interview, where they expressed that they liked typing information over choosing options from a drop-down menu because the drop-down menus were extensive and did not always have what they were looking for. The higher mouse movement could have resulted from the greater distance from the save and add buttons on the allergy page. Most significant differences for the other individual tasks were between HealthSpek and HealthVault, which have similar results when comparing the PHR systems.

B. <u>Discussion of PHR Systems</u>

Results on significant differences for dependent measures are shown in Table 3. These results are mostly consistent with the interview results, which will be discussed in further detail later. Another metric of interest are the omission errors, where Microsoft HealthVault errors were significantly higher than Health Companion and higher than HealthSpek. These errors mean that the participant thought the task was complete but did not fully enter the information provided. This is largely due to the separation of the blood pressure and weight tasks, as previously mentioned. The high omission error rate of Microsoft HealthVault could have influenced the total task time of the site or the participants' view of its usability. However, HealthVault also had the smallest standard deviation, which helps alleviate these concerns.

As expected, the order that the systems were used did have a significant effect on several factors, including task time and overall satisfaction. This was expected, because although the PHRs differed in their layouts, entering the same information three times allows for some learnability. However, since this was expected, the presentation order was altered for each participant. In addition, order was also used as a block to determine if means were still significantly different.

Some metrics that did not have a significant difference in means were interface quality and mouse clicks. Though information quality ratings were consistent with other CSUQ results in that Microsoft HealthVault scored the highest, these scores were closer to each other than the other metrics, with a difference of less than one for the highest-ranking PHR and the lowest scoring PHR (HealthSpek).

Though there was no significance in the total mouse clicks in the systems, the average number of clicks was directly related to the overall satisfaction for each PHR. For example, HealthVault had the highest average satisfaction with a rating of 5.57 and the average number of mouse clicks for the system was 227 clicks. On the other hand, HealthSpek scored the lowest on overall satisfaction with an average score of 4.36 and also had the lowest amount of average mouse clicks at 199 clicks. Initially, it was thought that a higher amount of mouse clicks would indicate a longer task time and user frustration, but these results do not support this assumption.

One reason that an increase in the number of clicks could be related to the overall satisfaction is the registration for Microsoft HealthVault. Several participants complained about the process, which had a sign up page that was difficult to find (see Figure 7), more fields for registration and including multiple verifications. Adding a new member was similar to the registration task, in that Microsoft HealthVault had more fields than the other systems. For these two tasks combined, HealthVault had 20 more clicks that HealthSpek. However, the task with the highest difference of clicks was in the allergy task, which had an average difference of 22 clicks between the two systems. The HealthSpek allergy entry only contained three fields where everything could be typed in. On the other hand, Microsoft HealthVault had more fields, including drop-downs that did not have the information that needed to be entered. This topic is discussed in more detail in the discussion of qualitative analysis.

One interesting, unplanned finding of the ANOVA testing was the significant difference between task time for males and females. No conclusive results on gender effects on task time have been found in the literature, though previous studies support the assertion that there may be a difference in how males and females interact with computer interfaces (Passig and Levin, 1995; Large, Beheshti, & Rahman, 2002; Lorigo et al., 2006).

C. <u>Discussion of Qualitative Interviews</u>

After the completion of all tasks in each PHR and the surveys, interview questions were asked of each participant. In addition, two of the survey questions from the CSUQ were open-ended questions: Please list positive and negative aspects of the system. Based on these subjective criteria, further analysis on the PHRs was completed. Each individual

PHR had several aspects that worked well and several that could be improved for improved usability in future.

More than half of the participants (53%) identified Microsoft HealthVault as their favorite of the three PHRs. Comments from the participants included the fact that the PHR layout was similar to those of other Microsoft products. One participant said, “Easy to use if familiar with Microsoft accounts or Windows Live.” Ten of 18 participants made a positive comment in the survey about the layout. Other participants also commented on the ease of adding additional records. Almost all of the users identified as Windows/Microsoft users. One participant that did recognize herself as an Apple user took over three minutes longer on HealthVault than the other two systems. It should be noted that this participant used HealthVault first, and that there was a significant difference in the order that participants used the systems. However, this would be an interesting topic for future research to determine if there is a significant difference in task time for Microsoft HealthVault based on whether participants identify as Windows or Apple users.

Negative comments were mainly focused on the lack of options on the drop-down menu for allergy reactions and the difficulty of registering for the site. The drop-down menu did not have all of the options that were on the information to enter, resulting in a variety of solutions from the participants. Some participants picked a reaction that was closest to the one described, while others selected the “Other” option and then typed the actual symptoms in the comments box. The registration screen also caused problems because the actual registration link was difficult to find. It also required signing up for a Microsoft account, which was a more lengthy process than other registration sites. The registration link is shown below (Figure 7), circled in red:

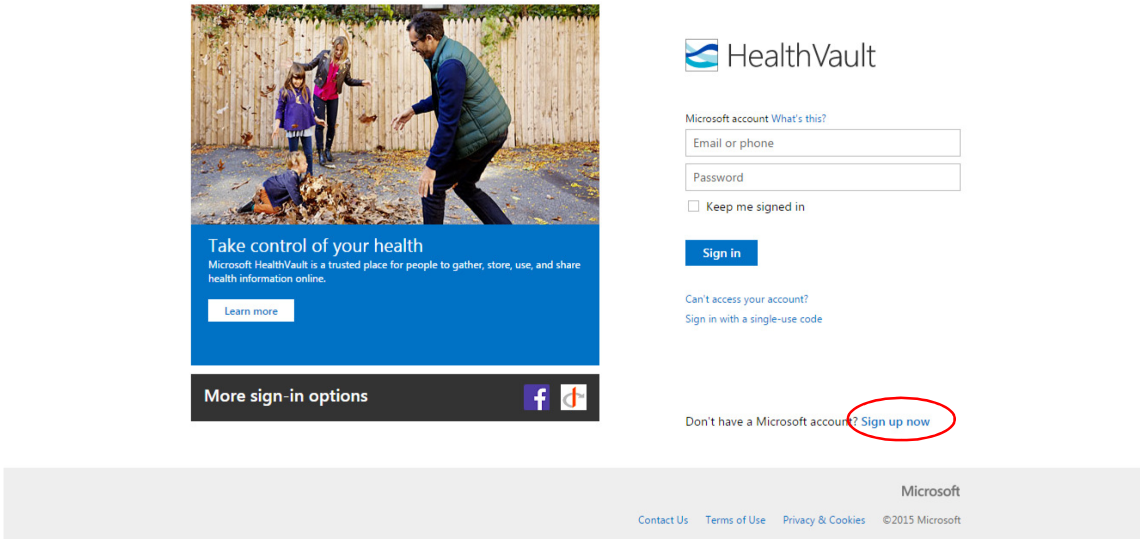

FIGURE 7 – Microsoft HealthVault Registration Page

The second best PHR, as picked by the participants, was Health Companion, with 31% of participants choosing it as their favorite. The reviews were mixed on the interface design, with nine participants saying that they liked the layout, and six participants saying that they did not like the interface. One participant stated, "Finding where to go to input information for the PHR and the modules constantly changing position when trying to navigate between them made the system difficult to use."

Others liked that all of the modules could be found on one page. One of the most difficult parts of the tasks was finding the link to the modules on the initial home screen. Images of the initial home screen and the modules page are shown below. Other positive comments included the ease of entering information and setting up an account. Negative feedback on the system included difficulty entering medication frequency and allergy reactions because both did not have options equivalent to the information the participants were supposed to enter. The location of the link to the modules for Health Companion is shown in Figure 8. Figure 9 shows the module page that the link takes the user to.

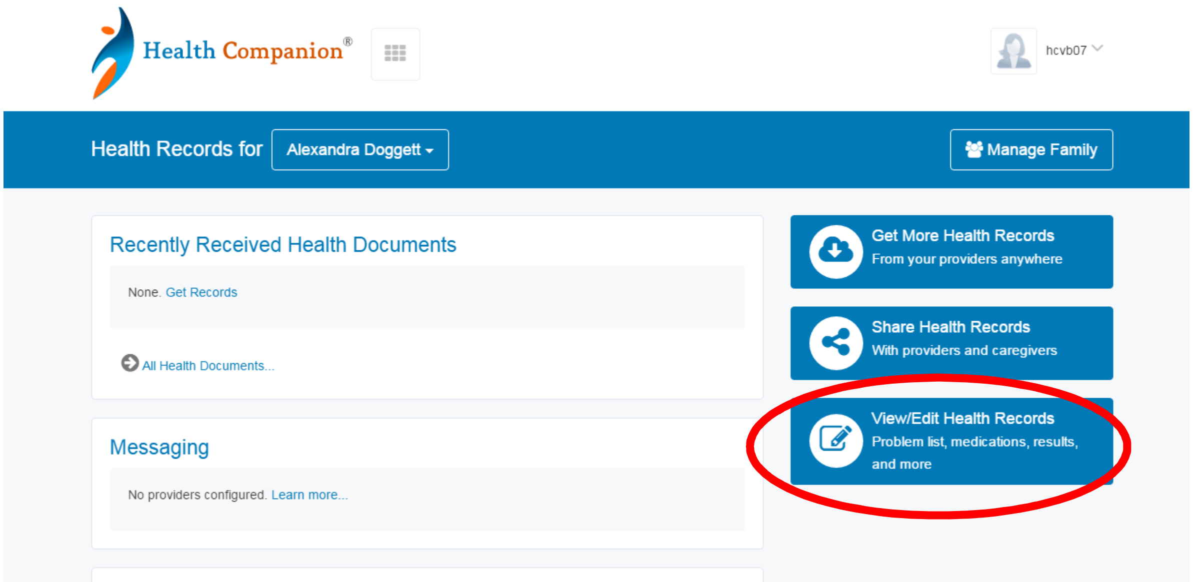

FIGURE 8 – Location of Modules on Home Screen of Health Companion

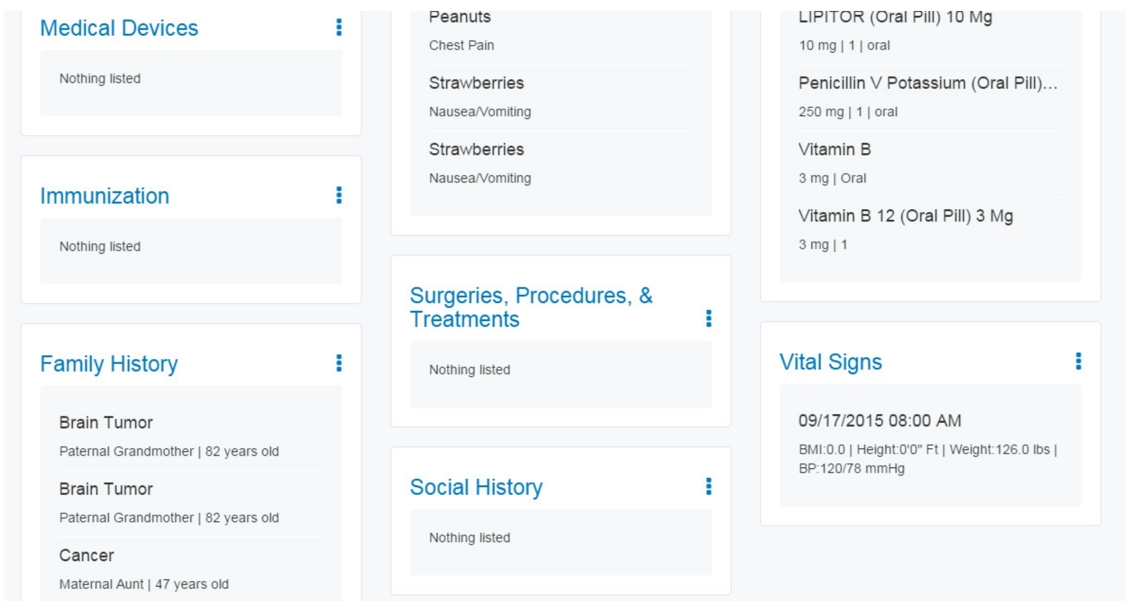

FIGURE 9 – Module Page of Health Companion

Finally, HealthSpek was the least popular of the three systems, with only 11% of participants identifying it as their favorite. This metric is interesting, considering that HealthSpek had lower overall task time than Health Companion. One of the most common negative comment from the participants about the PHR was entering information about blood pressure and weight. When entering this information, a screen pops up with average blood pressure or weight, and then the user can scroll through to get the desired blood pressure or weight. The main frustration came from the fact that the participants could not type or navigate to the number they wanted without having to scroll through all the numbers. This was especially frustrating for weight, where the participants were required to scroll through over 20 numbers three separate times.

Although multiple participants disliked the layout, most agreed that once they found where to enter the information, it was relatively easy to use the system. One participant said, “It was difficult to find the pages that I needed at first. The links on the left hand of the page were not very clear and didn't always have directly what I needed.” Another said, “User friendly layout once you know where to go for information.” Participants also liked being able to type information directly. Another issue that several participants mentioned was finding how to log out of the system. HealthSpek requires users to hit a plus sign button in order for the log out link to appear. This is shown in Figure 10. Figure 11 shows the log out location for HealthSpek, another issue that participants indicated.

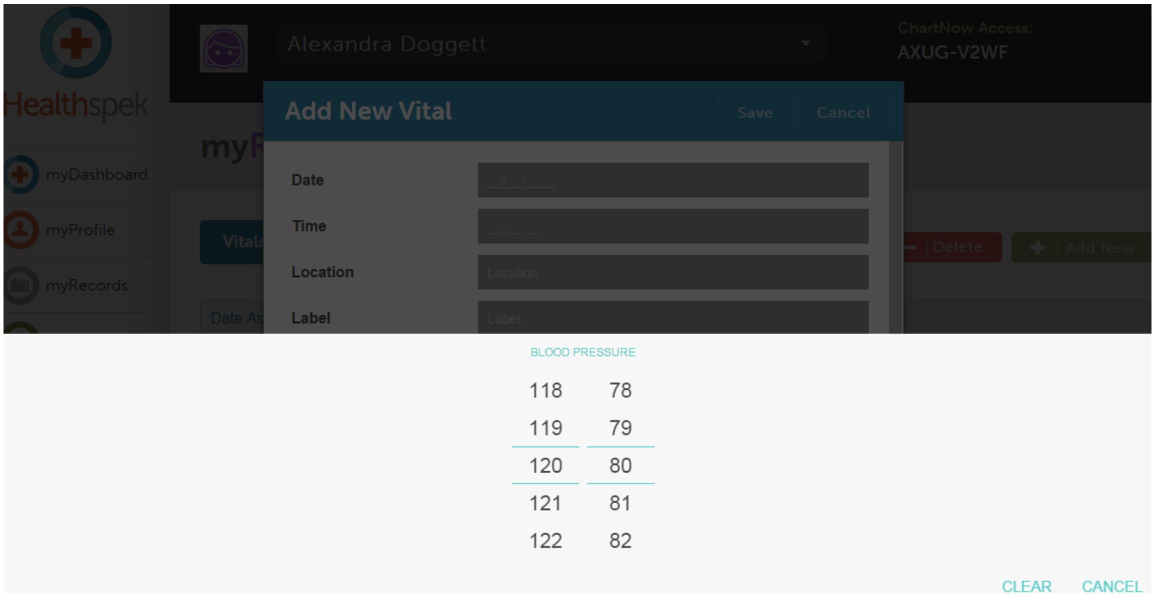

FIGURE 10 – Entering Blood Pressure in HealthSpek

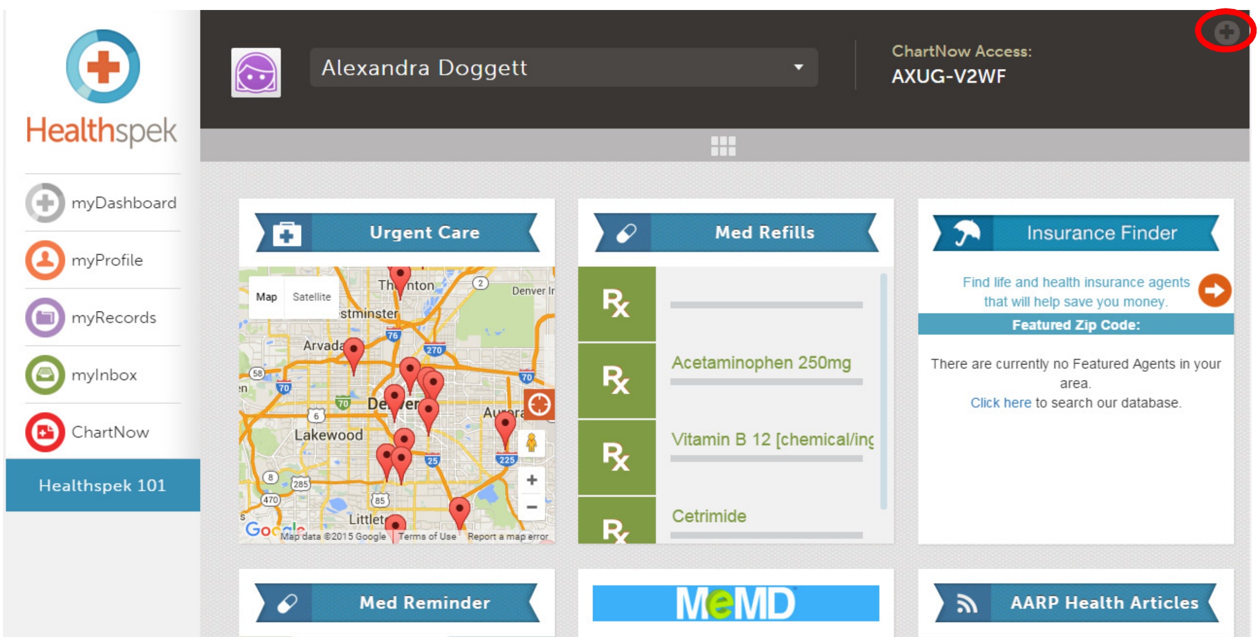

FIGURE 11 – Location of Log Out for HealthSpek

HealthSpek arguably has many similarities to the screen layout of iOS devices (see Figure 11). As pointed out by Norman and Tonazzini, Apple has taken an approach that places beauty and simplicity over usability and understandability (2015). Both the designs of Apple products and HealthSpek are based on colorful, simple applets that fill the screen. However, these applets negate some of the key principles of user-centered design, such as discoverability, defined as the ability to look at a system and discover all of the possible actions, recovery, defined as the ability to recover from errors, and consistency (Norman & Tonazzini, 2015). Many of these same traits are also missing from HealthSpek. For example, the location of areas to enter information differ. One can be entered in the main screen, but the rest of the modules are found through a side menu.

Other questions asked by the data collector were related to how participants would use PHRs in the future. Overall, many participants did not see themselves currently or consistently using a PHR to track their own health, mainly because they do not currently have any health issues. However, many concluded that if they had more health issues, they would be more likely to use a PHR system. Being able to update health records via an app or connect with other health information, such as exercise, would also be beneficial according to the participants. Most also said that they would be more likely to use a PHR for an elderly relative. In general, participants stated that they had the most trouble entering medication, stating that there were many fields and they were not as familiar with the terms.

VII. CONCLUSIONS

While the initial purpose of this project was comparative usability testing of PHRs for college-aged students, the study provided other insights as well. Similar to other usability studies found in the literature review, the study used multiple methods, including objective task metrics, a survey, and an interview to solicit feedback on the systems. This project provides a new addition to the literature in that it analyzes the usability of a system with a new user group and completed a comparative analysis of three leading Web-based PHRs. The initial hypothesis that there would be a significant difference in usability for each of the dependent measures and that Microsoft HealthVault would have superior usability based on these measures was partially supported by the results. While not all of the criteria had statistically significant results for the three different systems, such as task time, mouse movement, mouse clicks, and interface quality, many of the measures did have significant differences in their means.

Based on these results, it can be concluded that one system, Microsoft HealthVault, does have superior design and usability. HealthVault scored the best in all categories of the CSUQ, and in mouse movement, task time, commission errors, and keystrokes. HealthVault also had the best total task time, though it was very similar to HealthSpek (less than one second), and the results were not significant. For the CSUQ categories, HealthVault scored above 5 in all statistically different categories, while both Health Companion and HealthSpek were below 5 on a 7-point scale. Though more omission errors occurred in HealthVault than either other system, these errors were mainly induced from the similarities of the other systems and the separation of the blood pressure and weight information. In addition, the qualitative survey results also support the conclusion that

HealthVault has superior usability. The fact that these two separate sets of metrics converge into one conclusion further support the assertion that HealthVault has superior usability. One key aspect believed to be partially responsible for the ease of use of HealthVault is the design redundancy of the system. For example, most modules can be found in three different places on the home screen: in the toolbar at the top via the add button, on the menu on the left hand side under the appropriate category, or on the main page by scrolling down. This redundancy makes it easier for the user to find what they are looking for. Redundancy is also not as prevalent in the other two PHRs, where only certain buttons link to the information needed.

Though it can be concluded that HealthVault has a superior design from a usability perspective, it does not imply that the PHR is without error. There are still improvements that could be made to enhance the usability of the system. For example, several people commented on how they liked the interface of HealthSpek, which is more colorful than the HealthVault interface. Several participants also commented that some of the drop-downs limit the options and are not representative of the information they were trying to enter. Expanded drop-downs or typing options could be added to improve this. As found in the study by Fuji et al., there are still barriers to the adoption of Microsoft HealthVault (2015). Using the survey and open-ended survey results as a guide to improvement, Microsoft HealthVault has the potential to further improve its usability and distance itself from its competitors.

VIII. RECOMMENDATIONS

While the study was completed in a timely manner and met the objectives, there is still room for improvement and future research. If the experiment were to be completed in the future, several adjustments could be made. More participants could be included in the study to further verify the results. Random sampling would also be better to use instead of convenience sampling to make the results more generalizable. In addition, usability testing could be done on the individual systems to eliminate the order block. "Think-aloud" methodology could be incorporated into the experiment to gain further insights from the participants. Another improvement could be slightly altering the information to enter into each system, as Czaja et al. (2015) did, to decrease the effects of learnability for the participants as they re-enter information in each PHR. Also, as previously mentioned, another experiment could be conducted to determine if there is a significant difference in the time it takes participants to complete tasks based on whether they identify as Apple or Microsoft users.

APPENDIX I: IRB Outcome Letter

The Internal Review Board of the University of Louisville approved the experiment on the 21st of September, 2015. The full letter is disclosed on the following page.

UNIVERSITY OF
LOUISVILLE

Human Subjects Protection Program Office
MedCenter One – Suite 200
501 E. Broadway
Louisville, KY 40202-1798
Office: 502.852.5188 Fax: 502.852.2164

DATE:	September 21, 2015
TO:	Jason J Saleem
FROM:	The University of Louisville Institutional Review Board
IRB#:	15.0900
STUDY TITLE:	**A Comparative Usability Study of Web-based Personal Health Records**
REFERENCE #:	414961
DATE OF REVIEW:	09/20/2015
IRB STAFF CONTACT:	Name: Jacqueline S. Powell Phone: 852-4101 Email: jspowe01@Louisville.edu

This study was reviewed on 09/20/2015 and determined by the Chair of the Institutional Review Board that the study is exempt according to 45 CFR 46.101(b) under category 2: Educational tests un-linkable to individuals and no risks from disclosure .

Documents/Attachments reviewed and approved:

Submission Components			
IRB Study Application		Exempt	
Study Document			
Title	Version Number	Version Date	Outcome
Protocol	Version 2.0	09/16/2015	Approved
Questionnaire	Version 1.0	09/14/2015	Approved
Informed Consent	Version 1.0	09/14/2015	Approved

Please be advised that any study documents submitted with this protocol should be used in the form in which they were approved. Since this study is Exempt the consent form does not contain the IRB Approval Stamp.

Since this study has been approved under the exempt category indicated above, no additional reporting, such as submission of Progress Reports for continuation reviews, is needed. If your research focus or activities change, please submit an Amendment to the IRB for review to ensure that the indicated exempt category still applies. Best wishes for a successful study. Please send all inquiries to our office email address at hsppofc@louisville.edu

Thank you for your submission.

Sincerely,

Peter M. Quesada, Ph.D., Chair

APPENDIX II: Study Introduction & Information

Data Collector: Thank you for volunteering to help with this study. Today we will be asking you to complete a series of tasks on different Personal Health Record websites. I will be providing you instructions via written instructions and instructions given over the computer. Your workstation will be this computer. If you have any questions during the study, please use your best judgement and continue working. Before we get started, please take time to read the consent form and let me know if you have any questions before signing it.

This is your workstation. Once you have read over the instructions in front of you, feel free to begin. Once you begin, initial instructions will be shown on a pop up on the screen. Hit the start button to begin. Your first set of instructions will then appear on the screen. Once you have read them, you can hit start task to begin. The instructions will be hidden in a menu bar at the top of the screen. The menu bar contains two buttons: on the left will be end task, and on the right will be show instructions. If you need to check the instructions at any time, hit the show instructions button. Once you have completed the task, hit end task, and the instructions for the next task will appear.

Information given to the participants:

Have a seat at the computer. The screen in front of you should look like this (see next page):

Hit the red button to begin. After reading the instructions, you can hit Start Task to begin. If you have any questions about the information sheets or these instructions, please ask before you begin. When finished, hit end task and a new task shall appear. Due to time constraints, you may be asked to move on to the next task without completing all steps of a task.

Prompt: Background: You are a college student interested in better tracking your health. To do this, you have decided to use an online personal health record, or PHR, to keep track of your health. You have done some research and decided that three online PHRs seem to be the best fit for you. To determine which one will work the best for you, you have decided to register for each of them and enter some medical information on them. This way you can see which one is the easiest to use. Please do not actually enter any personal information into the PHRs. Use the information provided.

APPENDIX III: Subject Informed Consent Document

Subject Informed Consent

A Comparative Usability Study of Web-based Personal Health Records

Investigator(s) name & address: Jason J. Saleem, Ph.D.
Department of Industrial Engineering,
J.B. Speed School of Engineering,
University of Louisville,
Louisville, KY, USA, 40292

Site(s) where study is to be conducted: Center for Ergonomics, Room 303 Lutz Hall

Phone number for subjects to call for questions: (502) 852-2274

Introduction and Background Information

You are invited to participate in a research study. The study is being conducted by Dr. Jason J. Saleem, Mr. Dustin Weiler, and Ms. Alexandra Doggett. The study is sponsored by the University of Louisville, Department of Industrial Engineering**.** The study will take place at the Center of Ergonomics in Lutz Hall, room 303. Approximately 18 subjects will be invited to participate.

Purpose

The purpose of this study is to compare the usability of three online Personal Health Records (PHRs). According to Ozok et al (2014) PHRs are electronic applications for individuals to access, manage and share their health information in a secure environment. Specifically, we will investigate the time it takes to complete certain tasks on each site as well as the keystrokes to determine if one PHR is more usable than the others are.

Procedures
In this study, you will be asked to complete a series of scenarios and tasks in three different PHRs on the computer. For all of these tasks, you will be given fictitious health information to enter. All of the information that you need will be provided to you either through written instructions or via the computer. You will be given the same information and similar tasks for all three PHRs. Morae software will be active on the computer during the entirety of the study. This software will use the webcam to record your facial expressions as well as the screen so that mouse clicks and time to complete tasks may be reviewed.

Potential Risks
There are no foreseeable risks associated with this study. You will be asked to sit at the computer and complete tasks on the computer for up to an hour.

Benefits
A possible benefit of this study is that the finding may determine that one PHR is more usable than the other. In addition, it may help you better understand PHRs, determine which PHR is the best fit for you, and help you track your health in the future. The information collected may not benefit you directly. The information learned in this study may be helpful to others.

Compensation
You will not be compensated monetarily for your time, inconvenience, or expenses while you are in this study. As a token of appreciation for your participation, you will receive a t-shirt with the Center for Ergonomics logo immediately after the session.

Confidentiality
Total privacy cannot be guaranteed. Your privacy will be protected to the extent permitted by law. If the results from this study are published, your name will not be made public. While unlikely, the following may look at the study records:

> The University of Louisville Institutional Review Board and Human Subjects Protection Program Office.

All data collected will be secured in a locked cabinet and/or kept in a password protected computer. No identifiers will be kept.

Voluntary Participation
Taking part in this study is voluntary. You may choose not to take part at all. If you decide to be in this study you may stop taking part at any time. If you decide not to be in this study or if you stop taking part at any time, you will not lose any benefits for which you may qualify.

Research Subject's Rights, Questions, Concerns, and Complaints
If you have any concerns or complaints about the study or the study staff, you have three options.

> You may contact the principal investigator at (502) 852-2274.

If you have any questions about your rights as a study subject, questions, concerns or complaints, you may call the Human Subjects Protection Program Office (HSPPO) (502) 852-5188. You may discuss any questions about your rights as a subject, in secret, with a member of the Institutional Review Board (IRB) or the HSPPO staff. The IRB is an independent committee composed of members of the University community, staff of the institutions, as well as lay members of the community not connected with these institutions. The IRB has reviewed this study.

If you want to speak to a person outside the University, you may call 1-877-852-1167. You will be given the chance to talk about any questions, concerns or complaints in secret. This is a 24 hour hot line answered by people who do not work at the University of Louisville.

This paper tells you what will happen during the study if you choose to take part. Your signature means that this study has been discussed with you, that your questions have been answered, and that you will take part in the study. This informed consent document is not a contract. You are not giving up any legal rights by signing this informed consent document. You will be given a signed copy of this paper to keep for your records.

____

Signature of Subject/Legal Representative Date Signed

____ ____________________

Signature of Person Explaining the Consent Form Date Signed
(if other than the Investigator)

_______________________________________ ____________________

Signature of Investigator Date Signed

LIST OF INVESTIGATORS	PHONE NUMBERS
Jason J. Saleem, Ph.D.	(502) 852-2274
Alexandra Doggett	(859) 663-8527
Dustin Weiler	(502) 852-2274

APPENDIX IV: Provided Information

Provided Information

Information About Your Friend's PHR (if it auto logs out):	
Health Vault Email	hcvb07
Health Vault Password	PLEASEwork
HealthSpek Username	amdogg01
HealthSpek Password	PLEASEwork
Health Companion Username	hcvb07
Health Companion Password	PLEASEwork

Information for Adding a Person to Your Friend's PHR:	
Name	Changed each time
Sex	Male
Country	U.S.
Marital Status	Single
Birthdate	5/11/1992
Relation	Other

Information For Creating Your Own PHR:	
Name	Changed each time
Sex	Male
Country	U.S.
Marital Status	Single
Birthdate	5/11/1992
Email	Changed each time

Email Password	Changed each time
PHR Password	**PLEASEwork**
Phone	859-663-8527
Username	
** if doesn't work, add another 0	
Security Questions:	
What is your mother's maiden name?	Smith
What is your oldest aunt's name?	Sue
What was the first car you drove or owned?	Honda Accord

DO NOT SHARE ACCOUNT ACCESS

Information to enter into the PHRs (PLEASE ENTER ALL INFORMATION PROVIDED):

Allergies			
Name:	**Reaction:**	**Status:**	**Symptoms Began:**
Peanuts	Chest Tightness, Swelling	Active	1/1/2005
Penicillin	inflammation	Active	8/15/2000
Strawberries	nausea, vomiting	Active	4/1/1994

Blood Pressure & Weight				
Date:	**Time:**	**BP SYS:**	**BP DIA:**	**Weight:**
9/11/2015	9:00 AM	123	76	125
9/12/2015	9:00 AM	118	82	123
9/13/2015	9:00 AM	130	77	124

Family History		
Relation:	**Condition:**	**Age:**
Maternal Grandfather	Heart Attack	75
Paternal Grandmother	Brain Tumor	82
Maternal Aunt	Cancer	47

Medication			
Name:	**Dosage:**	**Frequency:**	**Type:**
Acetaminophen	1 pill (250 mg)	1 a day	Over the Counter
Vitamin B	1 pill (3 mg)	1 a week	Over the Counter
Cetrimide	Topical, 20%	1 a month	Prescribed

APPENDIX V

Interview Questions

Now that you have completed the study, we have a few questions about your experience:

Did one of the PHRs stand out to you as your favorite? Why?

Did you prefer more or less detailed requirements for entering information?

What was the hardest information to enter? On which PHR? Why?

What was the easiest information to enter? On which PHR? Why?

Do you see yourself regularly using a PHR? Why or why not?

Potential follow- prompts:

- What would motivate a student to want to use a PHR?
- Would they really use it on a computer – maybe they would use it if it were an app on a mobile device/smartphone?
- Do they see a doctor regularly? – If so, print or share info from PHR app?
- Would they want to use one for an elderly relative?

Do you have any other comments about any of the PHRs or the study?

APPENDIX VI: CSUQ Survey

CSUQ: Computer System Usability Questionnaire

Please rate the usability of the system.

strongly
disagree ***strongly agree***

1.	Overall, I am satisfied with how easy it is to use this system.	1	2	3	4	5	6	7
2.	It was simple to use this system.	1	2	3	4	5	6	7
3.	I can effectively complete my work using this system.	1	2	3	4	5	6	7
4.	I am able to complete my work quickly using this system.	1	2	3	4	5	6	7
5.	I am able to efficiently complete my work using this system.	1	2	3	4	5	6	7
6.	I feel comfortable using this system.	1	2	3	4	5	6	7
7.	It was easy to learn to use this system.	1	2	3	4	5	6	7

8.	I believe I became productive quickly using this system.	1	2	3	4	5	6	7
9.	The system gives error messages that clearly tell me how to fix problems.	1	2	3	4	5	6	7
10.	Whenever I make a mistake using the system, I recover easily and quickly.	1	2	3	4	5	6	7
11.	The information (such as online help, on-screen messages, and other documentation) provided with this system is clear.	1	2	3	4	5	6	7
12.	It is easy to find the information I needed.	1	2	3	4	5	6	7
13.	The information provided for the system is easy to understand.	1	2	3	4	5	6	7
14.	The information is effective in helping me complete the tasks and scenarios.	1	2	3	4	5	6	7
15.	The organization of the information on the systems screens is clear.	1	2	3	4	5	6	7
16.	The interface of the system is pleasant.	1	2	3	4	5	6	7
17.	I like using the interface of this system.	1	2	3	4	5	6	7
18.	This system has all the functions and capabilities I expect it to have.	1	2	3	4	5	6	7

19.	Overall, I am satisfied with this system.	1	2	3	4	5	6	7

List the most **negative** aspects

1.

2.

3.

List the most **positive** aspects

1.

2.

3.

REFERENCES

Case Studies and Customer Stories for Morae. (n.d.). Retrieved October 5, 2015, from https://www.techsmith.com/morae-casestudy.html

Crouch, P. C. B., Rose, C. D., Johnson, M., & Janson, S. L. (2015). A pilot study to evaluate the magnitude of association of the use of electronic personal health records with patient activation and empowerment in HIV-infected veterans. *PeerJ*, *3*, e852.

Czaja, S. J., Zarcadoolas, C., Vaughon, W. L., Lee, C. C., Rockoff, M. L., & Levy, J. (2014). The Usability of Electronic Personal Health Record Systems for an Underserved Adult Population. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 0018720814549238.

Fuji, K. T., Abbott, A. A., & Galt, K. A. (2015). A Qualitative Study of How Patients with Type 2 Diabetes Use an Electronic Stand-Alone Personal Health Record. *Telemedicine and e-Health*, *21*(4), 296-300.

Gee, P. M., Paterniti, D. A., Ward, D., & Soederberg, M. L. (2015). E-Patients Perceptions of Using Personal Health Records for Self-management Support of Chronic Illness. *Computers, informatics, nursing: CIN*.

Haggstrom, D. A., Saleem, J. J., Russ, A. L., Jones, J., Russell, S. A., & Chumbler, N. R. (2011). Lessons learned from usability testing of a personal health record. *Journal of the American Medical Informatics Association*, 18 Suppl 1, i13-i17.

Irizarry, T., Dabbs, A. D., & Curran, C. R. (2015). Patient Portals and Patient Engagement: A State of the Science Review. *Journal of medical Internet research*, *17*(6).

Kaelber, D., & Pan, E. C. (2008). The value of personal health record (PHR) systems. In *AMIA Annual Symposium Proceedings* (Vol. 2008, p. 343). American Medical Informatics Association.

Kannry, J., Beuria, P., Wang, E., & Nissim, J. (2012). Personal health records: meaningful use, but for whom?. *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine*, *79*(5), 593-602.

Kushniruk, A. W., & Patel, V. L. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of biomedical informatics*, *37*(1), 56-76.

Large, A., Beheshti, J., & Rahman, T. (2002). Gender differences in collaborative web searching behavior: an elementary school study. *Information Processing & Management*, *38*(3), 427-443.

Lewis, J. R. (1995). IBM Computer Usability Satisfaction Questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction, 7*, 57-78.

Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., & Gay, G. (2006). The influence of task and gender on search and evaluation behavior using Google. *Information Processing & Management*, *42*(4), 1123-1131.

Otsuka, S. H., Tayal, N. H., Porter, K., Embi, P. J., & Beatty, S. J. (2013). Improving herpes zoster vaccination rates through use of a clinical pharmacist and a personal health record. *The American journal of medicine*, *126*(9), 832-e1.

Ozok, A. A., Wu, H., Garrido, M., Pronovost, P. J., & Gurses, A. P. (2014). Usability and perceived usefulness of personal health records for preventive health care: A case study focusing on patients' and primary care providers' perspectives. *Applied ergonomics*, *45*(3), 613-628.

Learn How to Use Morae. (n.d.). Retrieved October 5, 2015, from https://www.techsmith.com/morae-uses.html

Monkman, H., & Kushniruk, A. (2013). Applying Usability Methods to Identify Health Literacy Issues: An Example Using a Personal Health Record. *Enabling Health and Healthcare Through ICT: Available, Tailored, and Closer*, *183*, 179.

Morae from TechSmith. (n.d.). Retrieved October 5, 2015, from https://www.techsmith.com/morae.html

Morae Version History. (n.d.). Retrieved October 5, 2015, from https://www.techsmith.com/morae-version-history.html

Norman, D., & Tognazzini, B. (2015, November 10). How Apple Is Giving Design A Bad Name. Retrieved November 30, 2015, from http://www.fastcodesign.com/3053406/how-apple-is-giving-design-a-bad-name

Passig, D., & Levin, H. (2001). The interaction between gender, age, and multimedia interface design. *Education and Information Technologies*, *6*(4), 241-250.

Personal Health Working Group. (2003). The personal health working group final report. *Washington, DC: Connecting for Health: A Public-Private Collaborative*.

Pinciroli, F., & Pagliari, C. (2015). Understanding the evolving role of the Personal Health Record. *Computers in biology and medicine*, *59*, 160-163.

Price, M., Bellwood, P., & Davies, I. (2015). Using Usability Evaluation to Inform Alberta’s Personal Health Record Design. *Driving Quality in Informatics: Fulfilling the Promise*,*208*, 314.

Segall, N., Saville, J. G., L'Engle, P., Carlson, B., Wright, M. C., Schulman, K., & Tcheng, J. E. (2011). Usability evaluation of a personal health record. In *AMIA Annual Symposium Proceedings* (Vol. 2011, p. 1233). American Medical Informatics Association.

Sheehan, B., & Lucero, R. J. (2015). Initial Usability and Feasibility Evaluation of a Personal Health Record-Based Self-Management System for Older Adults. *eGEMs (Generating Evidence & Methods to improve patient outcomes)*, *3*(2), 3.

Sivaji, A., & Tzuaan, S. S. (2012, July). Website user experience (UX) testing tool development using Open Source Software (OSS). In Network of Ergonomics Societies Conference (SEANES), 2012 Southeast Asian (pp. 1-6). IEEE.

Shimada, S. L., Brandt, C. A., Feng, H., McInnes, D. K., Rao, S. R., Rothendler, J. A., ... & Houston, T. K. (2014). Personal Health Record Reach in the Veterans Health Administration: A Cross-Sectional Analysis. *Journal of medical Internet research*, *16*(12).

Spil, T., & Klein, R. (2015). The personal health future. *Health policy and technology*, *4*(2), 131-136.

Tang, P. C., Ash, J. S., Bates, D. W., Overhage, J. M., & Sands, D. Z. (2006). Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association*,*13*(2), 121-126.

Tenforde, M., Jain, A., & Hickner, J. (2011). The value of personal health records for chronic disease management: what do we know?. *Family Medicine-Kansas City*, *43*(5), 351.

Tullis, T. S., & Stetson, J. N. (2004, June). A comparison of questionnaires for assessing website usability. In *Usability Professional Association Conference*(pp. 1-12).

VITA

Alexandra Doggett is currently a graduate student at the University of Louisville seeking a Master of Engineering degree. She graduated from the University of Louisville summa cum laude with her Bachelor's Degree in Industrial Engineering in May 2015. While in school, she received both the Thomas L. Ward Scholarship Award and the Robert C. Ernst Award from the Speed School of Engineering and completed four co-ops with The Walt Disney Company. In her spare time, she enjoys reading, intramurals, crafting, cooking, and cheering on the Louisville Cardinals and Cincinnati Bengals.