A Usability Evaluation of READY: An Electronic Data Collection Tool to Assess Rheumatoid Arthritis

Jennifer Fox
Senior Nursing Student
College of Nursing
The Ohio State University
Class of 2015
Fox.822@osu.edu

Po-Yin Yen, PhD, RN
Research Assistant Professor
The Ohio State University
310-C Lincoln Tower
1800 Cannon Drive, Columbus, OH 43210
Po-Yin.Yen@osumc.edu

Acknowledgements: This project was supported by the National Institutions of Health (P60 AR064172) and Genentech. (Principal Investigators: Dr. Jeffery R. Curtis & Dr. Peter Embi)
Table of Contents

Abstract ................................................................................................................................. 3

Chapter I: Statement of the Problem .................................................................................. 4
  Introduction ....................................................................................................................... 4
  Patient Reported Outcomes ......................................................................................... 5
  Rheumatic Disease Activity (READY) ........................................................................... 6
  Purpose .......................................................................................................................... 7
  Research Questions ........................................................................................................ 7

Chapter II: Methodology .................................................................................................. 7
  Research Design .............................................................................................................. 7
  Setting and Sample ...................................................................................................... 8
  Data Collection and Instruments ............................................................................... 8
  Data Analysis .............................................................................................................. 11

Chapter III: Research Results ........................................................................................ 11
  Entering Tasks .............................................................................................................. 13
  Homunculus Man ....................................................................................................... 13
  Medication Management ............................................................................................ 14
  Risk and Benefit Visualization .................................................................................. 15
  System Navigation ...................................................................................................... 16
  Error Types ................................................................................................................ 16

Chapter IV: Discussion .................................................................................................... 18
  Physicians’ Feedback .................................................................................................. 18
  Limitations ................................................................................................................... 18
  Challenges of Using Think Aloud Protocol ............................................................... 19

Chapter IV: Conclusion and Recommendations ........................................................... 19

References ........................................................................................................................ 21
Abstract

The advancement of technology and increase use of mobile applications has impacted the healthcare system. This study evaluates the usability of an application called READY (RhEumAtic Disease ActivitY). RhEumAtic Disease ActivitY (READY) was designed for in-office use to allow Rheumatoid Arthritis patients and clinicians to input, track, and manage clinical outcomes, such as global pain, fatigue, and various disease activity indices. The study focuses on physicians’ interaction with READY. Think-aloud protocol was used to capture physicians’ feedback and provide potential solutions to improve READY. Although usability issues were found, READY received positive feedback in its ability to benefit both patients and physicians.
Chapter I: Statement of the Problem

Introduction

The use of technology has been rapidly increasing and has had a big impact on the healthcare system (Free et al., 2012). Mobile applications can wirelessly communicate to a patient’s electronic health record, offering users easy and convenient access to this information (Plaza et al., 2013). They offer the opportunity to improve the communication between healthcare providers and patients (Ozdalga et al., 2012). These applications for mobile devices such as smart phones and tablets provide healthcare professionals and patients with a growing number of specialized tools and resources (Patrick et al., 2008). These applications encourage and promote the user to engage in more healthy behavior and the clinician can deliver health interventions (Blake, 2008).

Usability evaluation is a systematic way of assessing to what degree the application is effective (i.e. how well the application fills the desired need of the user), efficient (i.e. how much effort or time is needed to perform a specific task), and satisfactory to the user (i.e. does the user favor this application) (Wood, 1998). Usability testing help determine whether a product works for its intended user in a laboratory setting or in a real world setting. Usability testing is evaluating a product by testing it with its intended user and involves evaluation of its design, implementation, and evaluation of the system (Kushniruk & Boycki, 2014). For example, in this project, usability testing involves an observer and the user (e.g. physician) while they interact with a mobile application system (e.g. a mobile application) and perform a specific task (e.g. entering medication dosage) (Neilson, 1994). By performing usability testing, it ensures
the uses of medical applications are adapted to the users, the task they are performing, and that there are no negative or harmful outcomes (Bastein, 2010).

If the usability of the mobile application does not work for its intended use, significant errors could lead to negative outcomes for the patient (Kushniruk et al, 2005). Usability testing is vital to designing all informatics applications (Wakefield et al., 2015) and is critical that important information is presented in an effective way (Jaspers, 2009).

To perform a usability study, a number of usability evaluation methods are available (Peute et al., 2015). However, some of these methods are limited by practicality, accessibility of required human resources, and time to perform the evaluation study (Peute et al., 2015). The most preferred method of data collection is using the think aloud protocol. It is a well-recognized method and can provide information about the cognitive processes of subjects pertaining to usability problems (Peute et al., 2015). Thoughts that may be verbalized are feelings of frustration, confusion, or even relief when a task was easy to perform.

**Patient Reported Outcomes**

Patient reported outcomes via technology “generate a color-coded report reflecting an individual’s patient’s survey results over time and efficient depicts how different symptom items change in relation to each other…” (Smith et al., 2013). Research has shown when it comes to treating arthritis patients early diagnosis and prompt aggressive treatment substantially improve patient functional outcome and morbidity (Davis & Matteson, 2012). Treatment decisions should be made based on quantitative assessments of the patient’s disease activity (Saag et al., 2008). For example, the Health Assessment Questionnaire (HAQ) is
a comprehensive outcome measurement that assesses the hierarchy of patient outcomes in mostly two domains: disability and discomfort and pain (Ramey et al., 1992). Disability is assessed by eight categories: dressing, arising, eating, walking, hygiene, reach, grip, and common activities. Discomfort is determined by the presence of pain and its severity, location, and time of the day (Ramey et al., 1992). The time frame for disability and discomfort is based on the past week. The questionnaire is a paper and pencil self-administered survey that is given in an out-patient clinic to each RA patient before their appointment (See Figure 1). While electronic health record system (EHR) was implemented in more and more hospitals, additional data entry is required to transfer the patient reported outcomes on paper into the EHR.

**Rheumatic Disease Activity (READY)**

The Ohio State University and the University of Alabama Birmingham received funding from the Agency for Healthcare Research and Quality (AHRQ) and together has developed an application for handheld devices for patients with rheumatoid arthritis. Rheumatic Disease Activity (READY) is designed for in-office use to allow patients and clinicians to manually input data through a set of questionnaires. The clinicians can track patient data and their clinical
outcomes. It enables collection, storage, and trending over time of information on validated disease measures such as global pain, fatigue, and various disease activity indices. The benefit of using READY is that the patient’s scores are automatically calculated and their disease activity is stored in a graph so a trend can be seen over time. More importantly, medications and their doses are recorded in the graph which allows easier analysis of whether or not certain medications are working.

**Purpose**

This purpose of the study was to evaluate the usability of READY to identify human-computer interaction issues from the physician’s perspective. We observed physicians’ interaction with READY.

**Research Questions**

Our research question is “what are the usability problems that physicians encounter while interacting with READY?”

**Chapter II: Methodology**

**Research Design**

This was a cross-sectional study conducted at The Ohio State University and the University of Alabama Birmingham. A mobile application called RhEumatoid Arthritis Disease activity (READY), was developed to assist patient reported outcomes collected electronically and facilitate efficient patient care as well as shared decision making. To evaluate the use
usability of READY, data was collected through usability testing using video analysis software, Morae. The pilot study was reviewed and approved by the Institutional Review Board.

**Setting and Sample**

A total of fourteen Rheumatologists participated in the study from OSU and UAB. The six Rheumatologists from OSU were recruited via email and had volunteered to participate in a 30 minute interview to complete 18 tasks while using READY in a laboratory setting at Ohio State University CarePoint East. Each physician uses the electronic health record daily and is very familiar with the use of a computer, but each one has their own level of experience using a touch screen. The physicians range from handwriting notes during the appointment while using the electronic health record as minimal as possible to using the electronic health record as their main tool for recording notes.

**Data Collection Procedures and Instruments**

Data was collected from two rheumatology clinic between May and July of 2014. Participants were asked to use READY on an iPad for a series of tasks. Each evaluation lasted approximately 30 minutes and was conducted on an iPad. Think aloud protocol is a method of data collection used by researchers to provide empirical and procedural information when working with tasks performed by participants (Khajouei, 2011). Participants verbalize their feelings and general comments to gain insight into the cognitive part of the task (Jaspers et al., 2004). During each evaluation, physician’s hand movements with the READY application and
their verbal responses to the qualitative tasks and questions asked were video recorded. The physicians had no prior knowledge of how to use READY.

To carry out this process, a tripod was set up in a conference room with the camera recording only the iPad, the user’s hands, and their voice. The think aloud protocol was used to gain more information on the positives and negatives of the application. The users were given a list of tasks to perform without any guidance on how to perform said task (See table 1). Once the tasks were completed, series follow-up questions were asked to provide more detail on how this application could be improved, what did the user like/dislike, and what was most/least helpful. Each physician's interview was video recorded with their interactions with the application.

Using the video recorder in addition to the think aloud protocol was helpful for the data analysis process since the video could be re-played to determine which areas of the application the user had the most problems with and to determine any common problems experienced.
<table>
<thead>
<tr>
<th>Task Number</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assess the trend of the RAPID3 score (time it takes physician to find graph)</td>
</tr>
<tr>
<td>2</td>
<td>Enter diagnosis as rheumatoid arthritis</td>
</tr>
<tr>
<td>3</td>
<td>Enter a tender right knee on the homunculus</td>
</tr>
<tr>
<td>4</td>
<td>Enter a swollen 4&lt;sup&gt;th&lt;/sup&gt; and 5&lt;sup&gt;th&lt;/sup&gt; interphalangeal joint on the right hand</td>
</tr>
<tr>
<td>5</td>
<td>Enter a tender 1&lt;sup&gt;st&lt;/sup&gt; metacarpal phalangeal joint on the left hand</td>
</tr>
<tr>
<td>6</td>
<td>Enter a swollen AND tender left knee</td>
</tr>
<tr>
<td>7</td>
<td>Change the tender joint count to 5</td>
</tr>
<tr>
<td>8</td>
<td>Change the swollen joint count to 4</td>
</tr>
<tr>
<td>9</td>
<td>Input ESR at ____</td>
</tr>
<tr>
<td>10</td>
<td>Input CRP at ____</td>
</tr>
<tr>
<td>11</td>
<td>Select the physician global assessment as ____</td>
</tr>
<tr>
<td>12</td>
<td>Indicate that the tender joint count and the patient global score accurately reflect only their RA</td>
</tr>
<tr>
<td>13</td>
<td>Indicate that this patient is appropriate for a tight-control treatment strategy</td>
</tr>
<tr>
<td>14</td>
<td>Indicate the patient is currently on 8 tabs of Methotrexate (20mg)</td>
</tr>
<tr>
<td>15</td>
<td>Indicate the patient will decrease the dose to 7 tabs starting today (17.5mg)</td>
</tr>
<tr>
<td>16</td>
<td>Indicate the patient will start Humira today</td>
</tr>
<tr>
<td>17</td>
<td>Show that the patient has an 88% chance of getting better</td>
</tr>
<tr>
<td>18</td>
<td>Show that the patient has a 4% chance of having a side effect</td>
</tr>
</tbody>
</table>
Data Analysis

The video and audio recordings of the participants as they interacted with the device were analyzed using a software program Morae Manager for usability measures using markers to identify difficulties and errors the participants had (Clarke et al., 2014). The audio and video recordings were conducted in three phases. The first step was to review the video session and label all errors that occurred. The second step was to categorize READY’s tasks based on their affiliation with certain sections of the application (i.e. entering tasks, homunculus, medications, visualization, and navigation). The third step was to categorize each of the errors by error type (i.e. navigational, operational, when the participant was prompted, and software problems.)

Chapter III: Results

The results of this study were based on eighteen tasks each of the physicians were required to complete. The tasks included: enter diagnosis of rheumatoid arthritis, tender right knee, swollen fourth and fifth interphalangeal joint on the right hand, a tender first metacarpal phalangeal joint on the left hand, swollen and tender left knee, changing the tender joint count to 5 and the swollen joint count to 4, input a value for the ESR and CRP, select the physician global assessment, indicate that the tender joint count and the patient global score accurately reflect only their RA, indicate that this patient is appropriate for a tight-control treatment strategy, indicate the patient is currently on 8 tabs of Methotrexate (20mg), indicate the patient will decrease the dose to 7 tabs starting today (17.5mg), indicate the patient will start Humira today, show that the patient has an 88% chance of getting better, and show that the patient has a 4% chance of having a side effect. We further categorized the tasks into five
categories including entering tasks, the homunculus man ("little man"), medication management, risk and benefit visualization, and system navigation.

Out of the fourteen physicians that were analyzed, each one had experienced at least one error for a total of 144 error markers shown by figure 1. When comparing the errors by affiliation, 13.8% were entering tasks errors, 8.3% involved the homunculus-translated as "little man," 24.3% were errors with medications, the highest had to do with visualization at 39.2% and system navigation was 15.3%.

![Total Number of Error Markers by Task](chart.png)

**Figure 1**: The number of total errors between the fourteen physicians. The higher the number means more errors occurred.
Entering tasks

The category of entering tasks indicates that the physician had to input a number into a value box. Entering tasks (13.8%) had to do with tasks two and seven through thirteen. An example of this error was inputting the units for CRP and ESR-values that relate to a patient’s blood work. (See figure 2). The CRP should be “added and auto-calculated” before the appointment so this section of READY assumes that the patient has had blood work done before the appointment. However, some patients don’t require blood work or get it drawn after the appointment. Therefore, the physicians are unsure of what value to put in since blood work was done beforehand. Also, the physicians would enter the value but would try to move on without selecting the units resulting in an error message. This frustrated some physicians because they thought the units should be automatic.

Homunculus man

The second category was the homunculus man (8.3%), which had the smallest percent error, involved tasks three through six. The homunculus man is an outline of a person with accentuated joints. This allows the physician to indicate where the patient is experiencing tender and swollen joints and can be tracked overtime. The physician can interact with the virtual homunculus man that allows the physicians to
visually indicate pain severity for specific joints. One of the issues was that the right and left sides of the homunculus were assumed that the patient was in front of the physician for assessment (See figure 3). The orientation of the visual is not the same as the physician’s so when asked to perform a task on the right side of the patient without a real patient in front of them, the physician would initially perform it on the left side on the image. A few of the physicians expressed improvements that could be made. One noticed that the homunculus man does not have feet but would be important to have since arthritis patients with pain in the toes or ankles. Another problem that arose was concerning the colors that allowed the user to indicate if the joint was tender, swollen, and tender and swollen. One of the physicians was color-blind so he had difficulty completing the tasks. He was able to fully complete what was asked of him but the future of READY should include possibly patterns to indicate the status of joints instead of colors.

*Medication management*

The third category was medication management and had to do with tasks fourteen through sixteen. Medication management is described as indicating what medication is being prescribed, continuing the medication and which dosage the patient is on. When entering medications, the majority of the problems encountered occurred when they did not know how to update the medication list. The physicians would either continue the dose or change their dose but then would automatically try to click out of the box. When this happened, the medication dosage did not
update, they have to click “Update” to see the change occur (See figure 4). A few of the physicians commented on how this area could be improved. One physician said, “It’s more than just switching (medications) there’s another part of why—it’s more of a thought process and that is what the note section is there for.” Another physician said, “There should be a warning if the wrong dosage [was] entered [or] critical information was missing.”

Risk and benefit visualization

The fourth category was visualization (tasks 17 and 18) of risks and benefits of the patient’s medical condition and the medications they are on. Visualization is presented by the use of green or red smiley faces. The use of visualization gives the patient an idea of the percentage of getting better or having a medication side effect. When asked to enter “show an 88% chance of getting better,” the physician intuitively tapped the faces on the screen, but the system was not designed to function this way. The physician has to enter the percent into the box value which correlates with the number of smiley faces (See figure 5). This was the same that occurred with the risk of having a side effect on a specific medication, the only difference is the red color. One of the physicians stated that he “[doesn’t] mind the smiley faces and all but [just does not] understand how you could put in a percentage number...and [say] there is a 20% chance of feeling better or a 21% chance of feeling worse. That I would not be too keen on doing. It is probably the worst
part for me…” Another physician stated, “The side effect you can still make the case because this medication has been known to show to have a 5% chance of dizziness.”

System navigation

Finally, the last error type was system navigation (task 1), this is where the physician can locate different parts of the application. One way this can be done is by swiping up at the bottom of the iPad to allow for the graph to be shown. Errors occurred due to the fact that the directions to get to the graph were too small and located in an inconvenient location (Figure 6). One physician stated, “[The instructions] are in a place that is out of the way and without knowing what to look I would not have seen it.” As a result of this problem, physicians are confused when attempting to perform the task. For the navigation between pages, one physician believed “swiping up and down is tricky especially with instructions not in bold or not very noticeable”.

Error Types

From all the errors found in each task category, the error type was further classified. The error types include operational error, software problem, misinterpretation of data, the physician expresses inability, and subject prompted. Operational error is defined as a user operates the system incorrectly. For example, to find the diagnosis page, the physician would click next instead of swiping upwards. Software problem is when the user’s progress is limited
by an error in the applications coding. Misinterpretation of the data is when the physician incorrectly completes the stated task. The physician expresses inability to complete a task when unsure of how to carry it out and if needed, told or prompted what to do next. The results concluded that the medication and visualization section of the application were more difficult to perform (Figure 7).

![Total Number of Marker Errors by Error Type](image)

**Figure 7.**
Chapter IV: Discussion

When comparing the errors by affiliation, 13.8% were entering tasks errors, 8.3% involved the homunculus, 24.3% were errors with medications, the highest had to do with visualization at 39.2% and the navigation was 15.3%.

When comparing by error types, 39.6% were operational errors, 9.0% occurred with software problems, 3.5% involved misinterpretation of data, 15.3% happened with expressing inability, physician prompted was 19.4% and navigational errors were 13.2%.

Physicians’ feedback

Although usability issues were found in the system, most problems were able to solved after improvement. Overall READY received positive feedback in its ability to be beneficial for both patients and physicians.

Physician 1: "I think it is easier, it logs it in a graphic friendly way, something that is easy to look at and easy to follow."
Physician 2: "if made available I would love to use the program..."
Physician 3: “It’s a lot quicker and keeps me from having to go back and logging things already entered in the computer.”
Physician 4: "I think it will improve communication between the physician and patient.”

Limitations

There were a few limitations when carrying out the study. One limitation of this review is a small sample size. When there is a larger sample, the standard deviation from the mean becomes smaller. Having fourteen physicians poses a risk for more error. Nevertheless, these physicians are specially trained in rheumatology and understand the patient’s disease activity. Secondly, the study was conducted in a laboratory setting, which does not take into account common distractions doctors may experience during a clinical encounter. There is also a time
constraint within the clinic appointment which may cause more errors when the physicians do not have time to think about the task ahead. However, this allowed for in-depth though processes to gain a better insight on the issues involving READY.

Challenges with using think aloud protocols

1. Interruptions – A challenge while conducting the think aloud protocol was having physicians still being on call and having to take a phone call during the middle of the interview. These interruptions resulted in the researcher having to frequently stop between think aloud sessions to allow for participants to perform their normal responsibilities.

2. Remembering to think aloud – Another challenge to efficiently collecting data through this method was that some participants found it difficult to remember to continue verbalizing their thoughts as they performed each task. The researcher had to keep reminding some participants to verbalize their thoughts out loud. One participant stated that although the think aloud protocol was not difficult it was difficult to think aloud especially during tasks that were more simple to perform (Odukoya & Chui, 2012).

Chapter V: Conclusion and Recommendations

The biggest issues involved with READY were restricted to operational errors for medication management and risk and benefit visualization. For future research, improvements to the application include information put into READY needs to directly sync into the patients’ records to avoid repetitive work for the physician. Finally, a navigation bar would be useful so that it allows for more flexible movement between tasks instead of simply the “Next” button.
Dissatisfaction with the connection between the application and the EHR may prevent the clinicians from using READY in the clinical setting. These results may be communicated with the makers of the application to improve the system. This study was able to identify usability problems that may be impeding the use of READY. Although usability issues were found, READY received positive feedback in its ability to benefit both patients and physicians.
References


