

Evaluating the Effect of Display Size on the Usability and the Perceptions of Safety of a
Mobile Handheld Application for Accessing Electronic Medical Records

by

Simon Minshall

Bachelor of Science, University of Westminster, UK, 1995

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the School of Health Information Science

© Simon Minshall, 2018
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy
or other means, without the permission of the author.

Supervisory Committee

Dr. Andre Kushniruk, Health Information Science, University of Victoria

Supervisor

Dr. Elizabeth Borycki, Health Information Science, University of Victoria

Departmental Member

Abstract

INTRODUCTION: While mobile device use by physicians increases, there is an increased risk that errors committed while using mobile devices can lead to harm. This mixed-method study evaluates the effects of screen size on clinical users' perceptions of medical application usability and safety when interfacing to critical patient information. In this research, two mobile devices are examined: iPhone® and the iPad®.

METHOD: Eleven physicians and one nurse practitioner participated in a chart-review simulation using an app that was an end-point to an electronic health record. Screen-recording, video-recording and a think-aloud protocol were used to gather data during the simulation. Additionally, participants completed Likert-based questionnaires and engaged in semi-structured interviews.

RESULTS: A total of 105 usability, usefulness and safety problems were recorded and analysed. A strong preference was found for the larger screen when reviewing patient data due to the large quantity of data and the increased display size. The smaller device was preferred due to the devices portability when participants needed to remain informed when they were away from the point of care.

CONCLUSION: There is an association between screen size and the perceived safety of the handheld device. The iPad was perceived to be safer to use in clinical practice. Participants preferred the iPad® because of the larger size, not because they thought it was safer or easier to use. The iPhone® was preferred for its portability and its usefulness was perceived to increase with greater distance from the point of care.

KEYWORDS: Clinical information systems; Electronic medical records; Physician satisfaction; Usability; Usefulness; Safety; Error; Testing; Mobile device; Screen size; Smartphone; Tablet

Table of Contents

Supervisory Committee.....	ii
Abstract	iii
Table of Contents.....	iv
List of Tables	vii
List of Figures.....	viii
Acknowledgements.....	ix
Chapter 1: Introduction	1
Nomenclature.....	2
User Skill and Experience	3
Different People Use the Same Mobile Devices in Different Ways.....	4
Mobile Devices and Errors.....	5
Mobile Apps for Accessing Patient Data	6
Form Factor and Error on Mobile Devices.....	7
Research Questions	8
Chapter 2: Literature Review.....	9
Overview	9
The Usefulness and Usability of the Mobile Device in Healthcare.....	9
Methods	10
Results	11
The Role of Error in the Safety of New Computing Devices in Healthcare	31
Methods	31
Results	32
Discussion.....	39
The Problems with Smartphone and Tablet Form Factors.....	47
Results	48
Discussion.....	56
Conclusion	58
Chapter 3: Research Questions.....	59
Is There a Correlation Between the Size of the Handheld Device’s Display and the Perceived Safety of the iPhone® and the iPad®?.....	59
What Are the Preferences for iPhone® and the iPad® Form Factor in Clinical Use? ...	60
What Types of Usability Problems do Users have with Smartphone and Tablet Interfaces to an Electronic Medical Record?.....	61
Chapter 4: Methods.....	63
Recruitment of Participants	65
Randomization of Participants.....	66
Materials	69
Cases.....	70
Procedure	72
Demographics, instruction and training.....	72
Post-case interview.....	73
Final interview.....	74
Data Collection	76
Analysis of the Questionnaires	76

Analysis of the Transcripts	76
Analysis of Data Processing	81
Summary	81
Ethics Approval	82
Chapter 5: Results	83
Introduction	83
Participant Demographics	83
Notes	84
The iPhone®	86
iPhone® usability problems	88
iPhone® usefulness problems	95
iPhone® error and safety problems	100
iPhone® comments	103
The iPad®	117
iPad® usability problems	120
iPad® usefulness problems	128
iPad® error and safety problems	133
iPad® usability comments	137
iPad® usefulness comments	142
iPad® error and safety comments	148
Likert Results: iPhone® and iPad® Compared	151
Likert Scores	151
Summarized Likert Scores	158
Chapter 6: Discussion	167
General Findings	167
Skill and knowledge	167
The doctor-patient experience	170
Display Size and Perceived Safety	171
Form Factor Preferences	176
Portability Preferences	177
Usability Problems	179
Imaging, reports and data	179
Controls	182
Navigation	183
Chapter 7: Conclusion	185
Findings	185
Contributions of this Research to the Body of Knowledge	186
Contributions of this Research to Health Informatics Education	187
Which is the most useful device?	190
Which size device is the safest for use with patients?	190
Which size screen should be procured for use in healthcare?	190
Limitations	190
Future Research	191
References	193
Appendix 1: Ethics Approval	207
Appendix 2: Invitation to Participate	208

Appendix 3: Participant Consent Form.....	210
Appendix 4: Candidate Form	214
Appendix 5: Participant Session Scripts	215
Appendix 6: Semi-structured interview guide.....	217
Appendix 7: Technique for wireless screen recording of the iPhone® and iPad®	218
Appendix 8: Session Roster.....	230
Appendix 9: Demographics Questionnaire	231
Appendix 10: Post-case interview and questionnaire	233
Appendix 11: Post-case interview and questionnaire	234
Appendix 12: Final interview and questionnaire.....	235
Appendix 13: Sample Coded Transcript	238

List of Tables

Table 1: The final block for groups of four participants	69
Table 2: Usability Coding Dictionary	80
Table 3: Usefulness Coding Dictionary	80
Table 4: Safety and Error Coding Dictionary.....	81
Table 5: Problem and comment scores by code for the iPhone®	87
Table 6: Problem and comment scores by code for the iPad®	119
Table 7: Summary of Likert responses including total scores.....	165
Table 8: Aggregate median, mean, and, mode response values of all Likert statements	166
Table 9: Preferences for all Likert Statements	166

List of Figures

Figure 1: A side-by-side scale comparison of the iPad® and the iPhone GUI.	62
Figure 2: The research session screen-recording configuration.	64
Figure 3: Randomizing Screen Sizes with Cases.....	67
Figure 4: Randomizing Over a Pair of Participants.....	67
Figure 5: A Randomized Block of Four Participants.....	68
Figure 6: Screenshot of patients A and B.....	71
Figure 7: Anatomy of a coded excerpt.....	78
Figure 8: Participants' Age Ranges.....	83
Figure 9: Participants' Experience.....	84
Figure 10: The iPhone® app main dashboard participants' view.....	86
Figure 11: iPhone® Orders View.....	89
Figure 12: Charts on the iPhone®.....	90
Figure 13: Edit mode on a record revealed.....	101
Figure 14: An example of a spider or stick chart as seen in the app.....	106
Figure 15: Percentage distribution of iPad® codes by category and type.....	117
Figure 16: The iPad® app showing the dashboard.....	118
Figure 17: Percentage distribution of iPad® Usability codes by category.....	120
Figure 18: The Uranalysis table as seen on the iPad®.....	123
Figure 19: Imaging report with an incorrect X-ray image.....	130
Figure 20: Page view showing a list of allergies centre-left.....	137
Figure 21: Likert result charts device perceived safety.....	152
Figure 22: Likert responses for remaining statements.....	155
Figure 23: Likert responses to statements concerning the app.....	156
Figure 24: Results for overall device preference.....	157
Figure 25: Likert responses for all participants' responses.....	158
Figure 26: Mean Likert Scores for each question.....	160
Figure 27: Median Likert Response for each question.....	161
Figure 28: Mode of Likert responses for each question.....	162
Figure 29: The Welcome Screen on the iPad®.....	221
Figure 30: The VitalChart main iPad® user interface.....	222
Figure 31: VitalChart's iPad® patient Dashboard.....	223
Figure 32: An example of a document on the VitalChart iPad® App.....	225
Figure 33: First three screens on VitalChart's iPhone® App.....	226
Figure 34: Patient data navigation on the VitalChart iPhone® App.....	227
Figure 35: Drilling down to details on the VitalChart iPhone® App.....	228
Figure 36: Navigating and viewing clinical documents on the VitalChart iPhone® App.....	229

Acknowledgements

I thank my family, Margaret, Nate, Desmond, Elijah, Zane, Judith, Ruth and John, for their constant encouragement and support.

Thanks also go to all the members of staff in the Department of Health Informatics at the University of Victoria, particularly Professors Andre Kushniruk and Elizabeth Borycki whose advice and insight was invaluable and necessary for this thesis to exist.

TA special thank-you to my colleagues in the Department who have helped to pull me through the more difficult times.

Chapter 1: Introduction

This study evaluates the effects of mobile device screen sizes on clinical users' perceptions of medical applications usability and safety, when interfacing to critical patient information. In this research, two popular mobile devices are examined: the iPhone[®], a pocket-sized mobile device, and the iPad[®], a larger tablet computer the size of a small magazine. These devices are mobile computers that provide an interactive experience to the user by graphically displaying a colourful user interface (UI); they respond to the user's touch with a variety of gestural actions such as taps, swipes, and presses. They also contain sensors and components which allow them to sense: motion via an accelerometer, location via a global positioning system, and elevation via a built-in altimeter. Other features of these devices include a high-resolution camera, microphone, and a loudspeaker. Telecommunication and network access are achieved wirelessly.

These mobile devices are shipped with software application packages, referred to henceforth as apps, which provide basic functionality. Users can use them for web-browsing, communication, note-taking and scheduling. The iPhone[®] became available to the market in 2007 and was the first of the mobile devices to offer a gestural-input display and a marketplace for software apps named App Store.

The user can choose to extend these devices' functionality by downloading apps. For example, users may choose to monitor their physical activity by downloading a fitness-tracker app. Physicians may choose to download an app to retrieve medical records or an app that connects to an electronic medical record system (EMR).

The iPhone[®]'s form factor, its physical size and shape, allows it to fit into a trouser, or lab coat pocket. The iPad[®]'s form factor is too large for a standard pocket and needs to be carried by hand or by some other means.

App developers may optimize the user's experience of their app on each device. If they do so, then their users may have a different experience with the same app on each device. Even if developers do not optimize the user experience for each device, end users will necessarily have a different experience because the devices have different form factors. Given the different form factors and use cases, users' mistakes, errors, problems, and preferences may also vary (Alsos, Das, & Svanæs, 2012). Errors exist and can cause harm particularly if the device plays a role in patient care (Momtahan, Burns, Sherrard, Mesana, & Labinaz, 2007). In healthcare, the issues of usability and safety have come to the fore with the widespread deployment of electronic medical records. This study investigates how the form factor of these two devices affects physicians' perceptions of usability and safety.

Nomenclature

The acronym PDA – Personal Digital Assistant – stood for “Personal Data Assistant” a term coined by Apple in the early 1990s to characterize their early handheld devices (Isaacson, 2011). PDA became a general term used to identify palm-sized data-driven devices and in this paper is a specific term referring to the class of handheld devices that must synchronize to a server in order to update and/or transfer data to and from the device. PDAs usually have network connectivity that is limited to communication with server software; an example is the Palm Pilot family of devices (“PalmOne,” n.d.). In this review, this type of device is referred to as a PDA.

Products such as the iPhone[®], Android phones and late generation PDAs that do have wireless internet connectivity examples are the Windows phones, 2008 Palm Treo and the Samsung devices. These pocket-sized devices are referred to as smartphones.

Products that offer the same network connectivity as the small mobile devices, but are physically larger, are referred to as tablet computers. Examples are the iPad[®] and Samsung Galaxy products. The term handheld refers to either a mobile device or a PDA in cases where the distinction is not relevant. The term PDA, smartphone, handheld or tablet are collectively and singularly synonymous with the term mobile device in this paper.

User Skill and Experience

The iPhone[®] became available in Canada in 2007. Seven years later, Catalyst, a Canadian technology market-research organization, found that 55% of Canadians used a phone-sized mobile device. This percentage rose to 68% in 2015 and 76% in 2017. Since their introduction, mobile devices have been continuously improved, with developers adding features, improving sensors, and increasing the complexity of software used by these devices. The manufacturer of the iPhone[®] and iPad[®], i.e., Apple, improved their products' usefulness by allowing third party software developers to create and sell applications, apps, for use on the devices ("8 Years of the iPhone[®]: An Interactive Timeline," 2014).

The complexity of a telephone from the 1980s compared with that of a mobile device of the 2010s would likely seem incredible if one were to predict that in the 1980s. Complex apps can be difficult to learn, and they require skill to use. The mobile device is not a telephone, it is a general-purpose computer. The telephone on a mobile device is

an app. A medical app running on a mobile device is another package of software. It would seem, even to an uninformed observer, that computer-literacy is a skill needed to successfully operate a mobile device.

Device manufacturers, such as Apple, produce User Interface guidelines to aid software developers in the design of their apps (*iOS Human Interface Guidelines*, 2012). The guidelines specify rules to follow for the visual appearance and behaviour of all apps; the rationale being that by following the guidelines, users will have a similar experience on every app. In other words, if a user knows how to operate one app, then they should be capable of understanding other apps. Researchers studied a group of low-socioeconomic-status adults who lacked computer-literacy in order to discover whether this group could successfully navigate a mobile health app (Miller et al., 2017). The researchers found the only predictor for needing assistance they was a lack of experience with the devices. Although a learning curve exists with mobile devices, computer literacy was not required.

Different People Use the Same Mobile Devices in Different Ways

Consider that a computer workstation is used to gather data in a clinical setting. The same data may be reviewed later the same day with a tablet during a group discussion in another clinical setting and viewed again in the evening on a mobile device. Clinicians will interact with each device, i.e., the iPhone® and iPad® with their different form factors, in different ways depending on the users' context and preferences. Form factor may play a role in preferences; for example, the iPad® may be preferred during group discussions or patient encounters because of its larger screen size. The iPhone® may be preferred for a quick review while jotting a clinical note or scheduling reminder while standing in an elevator. By studying physicians in a simulated clinical setting, we can

learn what unique issues and problems are encountered through the use of these mobile devices.

Mobile Devices and Errors

Users often commit many errors using mobile devices. Frequently these are trivial mistakes with trivial consequences; for example, a misspelled a word in a text message. Errors while composing text messages can be more serious; for example, a car accident may be caused by a distracted driver composing a text message (Wilson & Stimpson, 2010). This is an example of a technology-induced error (Borycki, 2005). Although there are many distractions for vehicle operators, it is important to note that this particular attention-theft did not exist prior to the introduction of the mobile phone.

Technology-induced errors may also occur while a clinician uses a mobile device to access information or to place an order for treatment in a clinical setting. There may be potentially serious consequences and implications for safety when using mobile devices. Errors may be caused by interaction with the mobile device leading to a typo, an incorrect medication order, or due to the distracting effects of having to focus attention on the user interface rather than on the task at hand. Handheld mobile devices can be difficult to use accurately and concerns for their safe use in clinical environments have been raised in the literature (Horsky, Kuperman, & Patel, 2005).

New technology is known to cause new kinds of human error (Kushniruk, Borycki, Anderson, & Anderson, 2009). A variety of papers, published about the harmful effects of negative unintended consequences that arose as a result of rapid adoption of new mobile device technology, warn of the need for expedient usability testing to

counterbalance the detrimental effects of mobile device use in healthcare (Coiera, Ash, & Berg, 2016; Kushniruk, Nohr, & Borycki, 2016).

Mobile Apps for Accessing Patient Data

Many mobile utility apps are developed by third parties, and used by clinicians to perform a variety of tasks in healthcare such as: scheduling, reference, and calculation (Yaman et al., 2015). Another type of app that is available allows clinicians to view patient data stored within a hospital's EMR, an example of which is m-EMR, an app used in a 2700 bed tertiary hospital in Seoul, South Korea (Junetae Kim, Lee & Lim, 2017).

The main purpose of the m-EMR app is to allow clinicians to read patient information. It comprises four default menus and several submenus. The default menus provide patient lists, and users can choose one of the following menus: inpatient list, operation patient list, consult patient list, and emergency patient list. Once a patient is selected, their patient data is available to review. Orders are shown but cannot be created on the app and no data is actually stored within the device itself, all data is retrieved from a remote server (Kim, et al., 2017).

The m-EMR app is similar to the app used for the experimental part of this thesis. The topography, patient lists and the read-only nature all support the primary use case, i.e., the convenience of accessing patient data without requiring a fixed terminal nor a computer-on-wheels. This scheme would only be successful if most clinicians possess mobile devices; as discussed above this seems to increasingly be the case.

Form Factor and Error on Mobile Devices

Interactivity differs on devices of different size; for example, the iPhone® may be used single-handedly for data entry while the user is standing and holding the device because the display is sized in such a way that a typical user's thumb can touch any part of the UI. Interaction style and device size affect how a large or small device is used. For example, the two-handed style on a smartphone-sized mobile device is more difficult and less favoured on a tablet-sized device (Restyandito, 2017). The iPad®, cannot be used single-handedly while the user stands because the user's thumb cannot cover the whole screen. The iPad® trades the iPhone®'s compactness for the ability to show more information on the larger screen.

Low-Cost Rapid Usability Engineering is a usability testing method that leverages the portability of computing equipment and peripherals, i.e., laptops and cameras to enable the researcher to bring the usability testing session to the participant rather than have a bespoke usability lab (Kushniruk & Borycki, 2006). By utilizing the Software Usability Measurement Inventory combined with Low-cost Rapid Usability Engineering, the usability of a mobile device and app can be measured in a manner most convenient to the participant (Currie, 2005). The users' actions on each device can be analysed for errors made using screen-recording methods. The number of errors measured on the iPad® can be compared to the numbers of errors measured on the iPhone®. User interaction can be analyzed for patterns suitable for machine-capture for future study. If device size correlates with errors made, then measuring the users' device preference could predict error rates. If the clinician prefers a mobile device – iPhone® or iPad® – and if the

relative difference in error rates is known between differing devices, then clinician error rates could be estimated.

As healthcare applications such as Electronic Medical Records (EMRs) become ubiquitous and mobile handheld devices become more available with many different sizes, the question of what type of form factor is ideal for use in clinical settings has emerged.

Research Questions

In this thesis, the researcher aims to answer the following questions:

- Is there a correlation between the size of the handheld device's display and the perceived safety of the iPhone® and the iPad®?
- What are the preferences for iPhone® and the iPad® form factor in clinical use?
- What types of problems do users have with mobile device interfaces to an electronic medical record?

Chapter 2: Literature Review

Overview

The literature review is composed of three parts:

- A discussion of the usefulness and usability of mobile handheld computers in healthcare. This includes a review of mobile device use within the medical community using application software (apps) whose primary function is involved in reading, writing or capturing medical data. This work includes a review of the past six years of research in the area and also a review of the research on clinicians' access to medical records particularly those used in patient encounters.
- A review of the role of error in the safety of new computing technology in healthcare. The knowledge that new technologies can lead to new kinds of errors is used to examine the state of safety with respect to the integration of mobile devices in the clinical workspace.
- A review of the problems that exist with the use of the mobile device from a form factor perspective. This search is the widest ranging and examines the literature from published indexed sources about issues and problems encountered specifically comparing the results from publications on differing screen sizes.

The Usefulness and Usability of the Mobile Device in Healthcare

This part of the literature review is concerned with the usability and usefulness of the devices and their impact on safety and error-inducing characteristics in healthcare. It

captures a view of mobile device use within the medical community using application software whose primary function involved reading, writing or capturing medical data over the past six years of published research. This search encompasses all aspects of mobile device use and narrows down to reveal research on clinicians' access to medical records particularly those used during patient encounters. The intent of this literature review is to answer the following research question: What are the preferences for mobile device form factors in clinical use?

Methods

Search strategy. A search was conducted for papers on the clinical use of mobile devices specifically when device was used to access clinical database systems such as the Electronic Medical Record. The PubMed database was searched using a set of terms for the devices and the concept of usefulness to construct the following Boolean query:

(((((mobile OR handheld OR tablet) AND (device OR computer)) OR iPhone®) OR iPad®) OR smartphone) OR android) AND (((((errors) OR safety) OR human factors) OR usability) OR ease of use) OR usefulness).

Both specific and general terms for the devices were included as keywords to gather as many papers as possible. Limiting the search to “smartphone” and “tablet” did exclude relevant papers. The query was limited to English language papers published from 2007 to 2017. The year 2007 was chosen because it was the year the iPhone® was first available on the market.

Review of identified studies. Papers that met the inclusion criteria outlined earlier in this document were retrieved for full review. Papers were then examined for duplicate

results. The set of full-text papers that remained was reviewed in full. Relevant themes and findings were extracted from the included papers. These are grouped, presented and discussed in the results section of this review.

Results

Several major themes emerged after a review of the retrieved papers on handheld and/or tablet use in healthcare: (1) device use for imaging and sensors, (2) focus on human factors, (3) device use for data collection, (4) focus on decision support, and (5) device used to access other health systems.

Theme 1: devices used for imaging and sensors. Six papers in the result set were concerned with imaging and device-sensor use. A paper from Greece describes the development of a handheld application to become a terminal for a hospital Picture Archiving and Communication System (PACS) using a Digital Imaging and Communications in Medicine (DICOM) protocol (Ninos et al., 2010). The handheld device used an approach that retrieved the image and displayed it via an internal website. This method ensured interoperability of the PACS-DICOM system and provided a familiar experience for clinicians. Results varied by imaging modality; for example, images of thyroid ultrasounds were deemed to be of sufficient quality, but images of micro-calcification were difficult to interpret. Little consideration was given to the relatively poor display quality of the devices; no consideration was given to viewing, interpreting and diagnosing using the device with an uncalibrated display in environments with varying levels of lighting. If these devices were to be considered diagnostic tools for viewing imaging studies, then one would expect more attention to be paid to the quality of the viewing environment, ambient lighting, contrast ratios and other

critical image-assessment parameters. Given this, the mobile devices seemed to be assessed for their novelty value and not for their usefulness as diagnostic tools.

Conversely, an app used to visualise urinalysis results on a smartphone was described in a paper by Ra, Muhammad, Lim, Han, Jung, and Kim (2017). The researchers investigated the usefulness of using a smartphone as a medium for displaying colour swatches produced by lab tests. The findings show that the device and app delivered accurate results under various environmental illumination conditions without any calibration requirements (Ra et al., 2017).

Similar research examining image quality in cellular mobile device imaging from Korea reported that heavily compressed images transmitted over the cell-phone data network resulted in usable transmissions of CT images from a PACS to a clinician in the field (Dong Keun Kim, Kim, Yang, Lee, & Yoo, 2011). The system described was of similar visual quality, but more heavily compressed than the system described by Ninio (2010) which concluded that the imagery was not usable. Mobile display technology has greatly improved since the introduction of the iPhone® in 2007 with larger high-resolution displays. These tools have potential for usefulness.

Just such a device was described in Ramey, Fung, and Hassell's (2011) article in which the researchers developed a pathology application for remotely viewing frozen slides with an iPad® tablet computer. Ramey (2011) concluded that although quality and resolution are acceptable, the user interface of the system as a whole proved to be an obstacle to effective clinical use, i.e., poor usability reduced the systems acceptability.

Wound measurement was performed using disposable paper rulers. The paper was physically placed on the patient's body and the width and breadth of the wound was

measured by reading the rulers. Two papers, one from the United States (Sprigle, Nemeth, & Gajjala, 2012), the other from New Zealand (Hammond & Nixon, 2011), described using a handheld device to optically measure the dimensions of wounds. Both approaches used a handheld device mounted onto a separate case with the measuring optics providing photogrammetric data, enabling it to make an accurate 1:1 scale measurement. Both approaches seemed to use similar methods and it was unclear whether they were the same device or different implementations of similar ideas. The results showed that a device's mobility is of importance in the study; both papers also related that the two primary benefits were: (1) the accuracy and repeatability of measurement, (2) the measurement was done in a non-contact manner. This was an interesting example of healthcare practitioners borrowing ideas and techniques from varied sources including remote sensing in this case.

In a paper from Oxford, England that studied Parkinson's disease (Joundi, Brittain, Jenkinson, Green, & Aziz, 2011), the authors reported on a novel approach of assessing and measuring the tremor experienced by patients. They did so by using the accelerometer in the iPhone®. This sensor measured changes in motion with high sensitivity. It was used by a free app, *iSeismometer* (Takeuchi & Kennelly, 2010), to record the tremor data of a patient with Parkinson's. The app performed frequency analysis on the retrieved motion data, and this capability was used in the study to identify the patient's dominant tremor frequency. The authors suggested that this device-software pair was likely the simplest and most cost-effective way to acquire a repeatable, accurate, automated measurement.

Theme 2: focus on human factors. Ten papers reported on human factors issues associated with clinical mobile devices. They covered the use of (C. A. Woods & Cumming, 2009) and the validation of (van Duinen, Rickelt, & Griez, 2008) electronic, visual analogue scales in GUI designs. The mobile devices were reported as being well suited to this type of user-interface control because the control needed little decoration, i.e., numbers, indicators and other graphical elements to inform the user. In Canada, Woods (2009), compared the results of paper versus tablet computers using the VAS and found no significant difference between the results of the two media. The researchers concluded that the choice of device did not affect the VAS results. In the Netherlands the VAS was effective on a tablet computer and was used as an instrument-specific electronic Visual Analog Scale for Anxiety. The researchers emphasised that the novel method would be preferred over a paper version of the work (van Duinen et al., 2008).

Researchers in the United States, i.e., Turner (2011), reported on the process used to test a prototype computer-adaptive, patient-reported outcome tool for gathering data in headache research. The researchers contracted out a heuristic evaluation to two unnamed usability experts and used the results to improve the usability of their prototype. The paper detailed the actual results received from the contractors, but not the methods of the heuristic evaluation of usability. The researchers demonstrated not only the specifics of their tool's improvements, but also the notion that usability evaluation could be done without having any specific, in-house experts by delegating to a third party in an effective manner (Turner-Bowker et al., 2011).

From Norway, a published, more general paper on usability testing paid attention to the software's context-of-use and how that related to its usability. They provided an example

of a handheld-based EHR application designed, usability-tested and intended for clinicians to review lab results. The authors asserted that clinicians should be used to test clinical applications (Svanæs, Das, & Alsos, 2008). The final two of the ten usability papers discussed mobile applications used in different contexts. In Fromme (2010), software for gathering patient-reported outcomes was tested in two different age groups of similar cancer patients. Elderly participants represented a less-computer literate group and they reported significantly lower ease-of-use scores than did the younger group. The researchers concluded that it was unrealistic to expect uniform ease-of-use scores in mixed age groups and that each group could have a separate metric for usefulness and acceptability – each measure-set valid within its own context (Fromme, Kenworthy-Heinige, & Hribar, 2010). From Japan, a paper reported on the usability of Health Information System software designed for a thin client-computing environment (TCC) used in a new operating context (Teramoto et al., 2010). In TCC, a server ran the complete application and the user used their local machine to view it in a client window. No software ran on the client machine and all user interface inputs and outputs were sent over the network. The focus of this paper was to evaluate the usability of the same system on a different client, a wireless tablet computer with a pen interface. This was a good example of usability changing in differing operating contexts, because the user-input devices on a standard system, keyboard and mouse, are less sensitive to the latency found in TCC systems. Added to this was more latency from a wireless network. The evaluation showed that latency was acceptable if it was below a certain threshold but, what was more problematic, was the delay-scattering, and the variance in latency over time that caused the app to speed up and slow down. The new context of use for the

system in this case, resulted in the creation of a new parameter, delay-scattering, not measured previously (Teramoto et al., 2010).

Research from China reported on a new system's construction using a systematic approach to design a PDA-based Nursing Information Systems, based on the literature on mobile human-computer interaction (Su & Liu, 2010). By applying concepts such as pictorial realism, iconic menus, shortcut bars and consistency, the team produced a set of human interface guidelines, developed an application, and put it to the test. The paper critiqued the process, but did not offer much in the way of conclusion other than to provide suggestions for further research. A paper from the United States advocated for user-centered design using the Pocket-PATH application development approach as an example (Dabbs et al., 2009). By involving the end-user in aspects of design the paper argued the likelihood increased that the application would achieve its intended goals.

A paper on human factors from Norway examined the effect that form-factor had on doctor-patient communication and paid particular attention to the body language involved in clinical encounters (Alsos et al., 2012). The main thesis of the work was that the introduction of a new technology on the ward or other medical setting could affect the non-verbal communication and body language between a doctor and patient. The researchers gave examples of non-verbal communication, such as the act of putting a clipboard away to indicate that the session was over; or, that writing notes on paper then looking up to indicate they are ready to proceed. These acts, often subconscious, occur during device- use, i.e., PDAs, laptops, tablets, computers-on-wheels. The devices could affect natural communication and therefore a patient's experience of the encounter; thus affecting, in turn, the perceived quality of care. When a clipboard, for example, was

allowed for a body language action such as clipping the pen to the board, Alsos (2012) explained that the clipboard had this 'affordance'. New devices similarly affected non-verbal communication and that this should be measured in usability evaluations. One example suggested choosing a device-size that fits in the doctor's pocket because that affords putting the PDA away to signify the end of a session. Another was to use a cover on a tablet so that it could be folded closed, and indicated the end of the note-taking session before the beginning of some other activity. The paper concluded that affordances should be incorporated into design and possible design guidelines for mobile point-of-care systems for improved doctor-patient communication were provided.

Theme 3: device use for data collection. Data collection was the subject of several papers. This broad group broke down into Ecological Momentary Assessments¹ (EMA), devices used for data collecting in longitudinal fieldwork, telemedicine, clinical data collection and five other papers generally concerned with handhelds used in data collection. Data collection in this context excluded the implementation of an EHR, or web-portal to one, for the purpose of recording doctor-patient encounters that were discussed separately below. These 37 papers were concerned with more generic and varied collection activities. A clear example of data collection was in public health studies in Fiji, where researchers used mobile devices as data-entry tools and reported an increased efficiency and a reduction in errors and labour when compared to the results of a parallel effort using paper-based instruments (Yu, de Courten, Pan, Galea, & Pryor, 2009). Similarly positive results were reported from South Africa where initial costs

¹ The Ecological momentary assessment (EMA) involves repeated sampling of subjects' current behaviors and experiences in real time, in subjects' natural environments (Shiffman, Stone, & Hufford, 2008)

were recouped with device re-use in subsequent studies (Seebregts et al., 2009), and in Peru where device use also resulted in reduced treatment delays (Blaya, Cohen, Rodríguez, Kim, & Fraser, 2009), and in Tanzania where handheld in-device data validation reduced data collection omissions (Thriemer et al., 2012). Contrarily, a study in Kenya reported high rates of missing data exacerbated by poor internal infrastructure, software design, usability and training. This same study also noted that despite the lack of usefulness for data collection, the use of the handheld devices had an unforeseen public-health benefit at the macro level by identifying those clinics under-testing their populations for tuberculosis (Auld et al., 2010).

Two studies in the United States also reported the positive results of using handhelds in a public health setting, but for different reasons. In Vinney (2012), the handheld devices were used by child patients with speech-language disorders to assess quality of life or patient-reported outcomes. “The percentage of children who made answering errors or omissions with paper and pencil was significantly greater than the percentage of children who made such errors using the device” (Vinney, Grade, & Connor, 2012). Similar, positive conclusions were found whereby the children in the study had no developmental disorders (Martin, Ariza, Thomson, & Binns, 2008). Another EMA study, involved automatically prompting the participants – older adults with limited computer skills – to log their physical activity exercise (Wolpin, Nguyen, Donesky-Cuenca, Carrieri-Kohlman, & Doorenbos, 2011). The researchers required accurate, timely logs of patients’ activity. Paper-based diaries had often resulted in ‘diary hoarding’ where the log would be filled-in just prior to meeting with practitioners, rather than just after the activity being logged. Previous research measured the compliance of paper-based

logging of pain at pre-specified times, was only 11% and logging using automated prompts from a PDA handheld device increased compliance to 94% (Stone, Shiffman, Schwartz, Broderick, & Hufford, 2003). Though much improved, the PDA device lacked the immediacy of real-time data possible with the mobile device. The results provided a more detailed view of compliance. It varied with the participants' computer skills, such that intermediately skilled participants were more compliant (i.e., 83%) than those with beginner skills (i.e., 16%). Weekday prompts were twice as likely to result in logging than weekends. In addition to providing data on when best to provide prompting, the authors also stressed the importance of usability testing using subjects that belonged in the same cohort as the intended clinical research. For example, their population had two groups; one group was younger and the other older. The recommendation was to perform separate usability tests on each group (Wolpin et al., 2011).

Four other papers discussed device use in EMA participant self-assessments. Conclusions were also positive with respect to the utility of handheld devices (Epstein & Preston, 2010; Hachizuka et al., 2010; Luckmann & Vidal, 2010; Shively et al., 2011). Two of these four studies, i.e., Shively (2011) and Epstein (2010) mentioned the handheld device as part of the method, but did not refer to it again in the discussion, which suggested a tacit acceptance with respect to the use of the device. This was similar to the way that one would not mention the utility of the pencil in a study done before handhelds were common – an indicator of the devices' maturity in healthcare population studies.

Six of the seven papers concerned with telemedicine in chronic disease management reported positive results. Research from Spain proved the feasibility of PDA devices for

patient-use in telemedical diabetes care, resulting in more accurate data capture. This had a high acceptability with the added feature that the mobile device itself was used to control and program third-party hardware such as insulin pumps (García-Sáez et al., 2009).

Research from Norway demonstrated that it was possible to use a handheld food-log app and a commercial blood-glucose meter to assemble a system that effectively helped diabetes patients manage their medications, diet, and physical activity; this resulted in what is effectively a behavioral intervention (Eirik Årsand, 2010). The half-year study noted that enabling the patient to self-manage food intake caused a change in their motivation. Similarly, from the United States, we learned that cell-phone use for adolescent diabetes management was feasible (Carroll, DiMeglio, Stein, & Marrero, 2011), and that data-entry compliance relating to meals rose from 43% to 58% when a handheld-based data collection method was integrated with behavioral intervention (group therapy) to increase skills in the participants' self-management of the disease (Sevick et al., 2008). The handheld itself was acknowledged to be part of the treatment because of its ability to provide instant feedback, which in turn, increased the participants' sense of mastery. Research from Japan proved the feasibility of home-based, handheld food logs (Tani et al., 2009), and from Sweden (Riazzoli et al., 2010) reported that handhelds enabled patient-reported data in a case where previously only clinician-reported data was feasible. Finally, a study from the United States on pain management reported no difference in the quality of data collected with handheld versus paper, but did note that the handheld was easier for their participants to use and the EMA

aspect of handheld data-collection allowed a faster response to changes in patients' conditions (Marceau, Link, Jamison, & Carolan, 2007).

Data collection with handhelds was not limited to field studies and telemedicine. Their use in collecting data was also reported within healthcare facilities. Scenarios ranged from the relatively simple point-and-click barcode capture (Akiyama, Koshio, & Kaihotsu, 2010; Hayden et al., 2008), to more complex observe-and-note capture and voluntary reporting of medication administration errors (Dollarhide, Rutledge, Weinger, & Dresselhaus, 2008; Westbrook & Woods, 2009). All four papers demonstrated the feasibility and effectiveness of handhelds use for these tasks. The largest group of in-hospital data collection papers (n=7), investigated the collection of patient data using a handheld-based rather than paper-based questionnaires. A Radiology department built on the success of switching imaging modalities from film to digital by switching to a paperless workflow (Robinson, DuVall, & Wiggins, 2008). They replaced paper questionnaires by creating a web-based forms that patients could complete on tablets. The study reported that students from the Biomedical Informatics department were used as test subjects during the development phase of the project. Results were positive and usability findings were fed back into the development cycle to further improve the tool. Despite the fact that the sample population of participants was not representative of typical patients on the ward, i.e., all were university students, none was ill or really needed to be present in the radiology department were it not for the study. They reported that the device was easy to use, navigate, and read. They also had usability problems with the radio buttons, checkboxes and the handwriting recognition interface (Robinson et al., 2008).

The issue of context-of-use was discussed in a paper from the United States by Hess (2008) with findings from a much larger sample of 10000 patients who had completed a more general primary care-based questionnaire (Hess, Santucci, McTigue, Fischer, & Kapoor, 2008). Although the majority of users (84%) reported no difficulty completing the tablet-based questionnaire, some did, and within their dataset was enough information to determine predictors about who would likely have more difficulty than most in completing the questionnaire. These predictor variables included ethnicity, educational level and certain co-morbidities. The paper concluded with a simple caution not to overlook a minority of users who found this technology difficult to use (Hess et al., 2008). Another paper from the United States compared the completeness of data collection in a paper-vs-PDA study and concluded that the handheld computers, although they did produce more complete data than the paper method, were not superior to the paper forms due to loss, theft and technical difficulties encountered with the PDAs that needed to be synchronized with a server to transfer data from the device to a central computer (Galliher et al., 2008). The authors acknowledged that the use of wireless, always-connected, mobile devices, combined with a web-based approach, would solve some of the data-collection problems they encountered. This appeared to be a case of unfortunate timing, because the next generation of devices, the mobile devices, did just that and were made available the same year, soon after the publication of these results. Another paper published the same year reports no such difficulties in their evaluation of PDA use in the ER (Rivera et al., 2008). In both Rivera (2008) and Galliher (2008) it was the clinicians who were using the devices. Both studies measured data-entry error, though the methods differed. The error rates were 0.2 errors per PDA form, 1.6 errors per

paper form in Rivera (2008), and 35% for paper, 3% for PDA in Galliher (2008).

Although the measurements in the two papers were not comparable, it was interesting to note the magnitude of the difference for paper vs PDA errors.

Research from Germany published in 2012 analyzed a tablet computer from a usability and economic point of view, and measured the cost of the device and software for administering patient questionnaires against the costs of using paper (Fritz, Balhorn, Riek, Breil, & Dugas, 2012). They found that the tablet was well received by both patients and clinicians. They also found that the cost of a paper-based system would equal the cost of the table-based system in fewer than seven months. There were too many variables for this conclusion to be generalized to other settings, but for the test site with its specific pre- and post-processing needs, the tablet computer was a cost-effective solution (Fritz et al., 2012).

Research from neighbouring Switzerland investigated the errors associated with patient data entry on both a PDA and a laptop computer in a quiet environment to determine which device was best suited for clinical research (G. Haller, Haller, Courvoisier, & Lovis, 2009). The researchers' findings indicated that handheld devices should be used with caution because they doubled the data entry time and increased the risk of typing errors during the data entry process.

Theme 4: focus on decision support. Three papers on Decision Support Software (DSS) examined the utility of and attitudes toward PDA devices (Johansson, Petersson, & Nilsson, 2011; Kuiper, 2008; Schnall, Velez, John, & Bakken, 2011). Of the three papers, two focused on reasoning and decision-making among nursing students. The first of the three papers, a psychometric evaluation, suggested in its findings, that a handheld

device and a 14-item self-administered scale developed by Ray (2006), to evaluate physicians attitudes, was also appropriate for measuring attitudes of nurses, towards nurse-related DSS software (Schnall, Velez, John, & Bakken, 2011).

The second paper described a study whereby nursing students who were given PDAs were compared to a similar group of students without PDAs, and assessed after 14 weeks to determine the differences in learning, clinical reasoning and higher order thinking after using a PDA (Kuiper, 2008). The PDAs themselves were used to gather data about their use and a battery of assessments, scales and worksheets; they concluded that their investigation strengthened the research supporting the use of PDA resources in nursing curricula. The third DSS paper was an in-depth, single case study that followed a nurse in Sweden and examined her day-to-day use of a PDA (Johansson, Petersson, & Nilsson, 2011). On the positive side, device-use resulted in stress reduction and increased organization and user efficiency. On the negative side, concerns were expressed about data security, patient perceptions of device use and a lack of integration with the EHR system. The built-in calendar application was used as an example of tool-use that led to increased organization. The paper concluded with a statement that the PDA could be useful in healthcare.

Four papers in the DSS group discussed the development and design of new software targeted at mobile device use. Development work at the University of Massachusetts on a medication dosing support system for HIV/AIDS care, was done using a multi-disciplinary team to design a mobile system targeted at the third world, where patient numbers are high and mobile networks proliferate. The team's goal was to provide access to clinical information in remote areas where access was lacking (Sadasivam,

Gathibandhe, Tanik, & Willig, 2010). In a joint project between National University of Singapore and the University of Warwick, the researchers developed an internet-based database of chemotherapy regimens and drug-drug interactions (DDI) (Yap, Chui, & Chan, 2011). The authors stated that powerful anti-cancer drugs (ACDs) can have significant toxic interactions with other ACDs and the lack of DDI databases and software in this area prompted the development of a new tool to fill the gap.

Researchers at the University of Toronto undertook a qualitative study involving a DSS tool (Kastner et al., 2010). The researchers used focus groups in their efforts to develop a DSS tool for osteoporosis disease management. The conceptual model for the patient-operated assessment tool was transformed into a functional prototype using the findings from the focus group studies. This design-by-committee approach was described in Kastner (2010) and included many illustrations of practical changes to the tool's user interface that were intended to encourage participants to respond to the assessment questions and to do so in an honest manner, so that the validity of the assessment would improve. Similarly, research at Columbia University School of Nursing in New York led to the development of a pediatric depression screening instrument in electronic form for a PDA. In this research, the investigators attempted to determine whether using a PDA, used as a decision support system, would improve screening effectiveness (John et al., 2007). A paper-based questionnaire-style screening instrument was used to create a Palm PDA application, which, in turn, was used by a team of 24 nursing students to screen 124 children for depression. Their findings revealed concerns about the device, specifically the clinician's use of the device in front of the patient or participant. The nurses believed that the device was a barrier to the nurse-patient encounter. In the paper's discussion

section, the researchers suggested the idea that the “barrier effect” could be mitigated if the users were more skilled, implying that the barrier to was not the device itself, but the user’s awkward use of, and fumbling with the device. The researchers reported that the effectiveness of device-use could be improved by further user experience and practice with the device and by pre-planning strategies such as entering the patients’ initial information before the actual encounter began. This finding was also published two years later in a paper by Dawson and Kushniruk (2009).

Five papers described the use of existing mobile devices and software in clinical settings, and one in a simulation setting. A team in the United States investigated the role of intuition and the use of a mobile device DSS tool by anesthesia practitioners (Coopmans & Biddle, 2008). The participants, a group of certified registered nurse anesthetists (CRNA), were divided into two groups and put through a series of two simulated events: (1) one group was instructed to proceed based on their own knowledge and experience, and (2) the other was told to do the same, and to use the PDA. The results were mixed. The PDA group took longer to detect adverse events in both series and took longer to treat the patient in the first simulation. They also took less time to determine the treatment for the patient during the second simulation. In general, it would seem that the PDA group was generally less effective, though the report offered no such specific conclusion. The report stated the case for the PDAs' potential to reduce error in complex scenarios and asserted the validity of the simulation method as a research tool (Coopmans & Biddle, 2008). In an exploration of PDA software, a team at the Ottawa Hospital, built a PDA application to replace a paper-based DSS in order to aid the university of Ottawa Heart Institute’s nursing coordinators, when they answered more

than 2000 annual cardiac-care, tele-triage calls from patients with queries, new symptoms and emergencies (Momtahan et al., 2007). Their study assessed the viability of the PDA as a DSS tool to facilitate the transfer of knowledge from a highly skilled group nearing retirement to new staff on the ward. The decision to create the software was made because little was available in the marketplace. The result was a viable, effective DSS tool with the limitations of slow hardware and a cumbersome data entry mode (i.e., a stylus input device with a simplified text entry graphical code known as 'graffiti'). A similar, possibly concurrent study, also examined performance using a tablet-PC and stylus input device. This resulted in improved DSS performance, an above-neutral participant satisfaction and similar limitations. Research from Sweden evaluated nurses' experience with a customized mobile DSS, LIFe-Reader, in a qualitative study by Johansson, Peter and Nilsson (2010). The device, a PDA with a built-in barcode reader, was used by home-visit practitioners to scan barcode labels on patients' medications to detect DDIs and other potential events based on evidence gathered by the device. The study analyzed interview transcriptions and measured opinions on prevention of drug-related injuries with respect to safety, usability and usefulness. The paper concluded that this specific device had good potential in a homecare setting and suggested that such technology may reduce medication DDIs if used regularly (Johansson et al., 2010).

Swiss researchers reported on the use of PDA-accessed DDI databases in an outpatient clinic where a database designed for patient safety was successfully used to identify potential DDIs (Dallenbach, Bovier, & Desmeules, 2007). Although the paper focused on the utility of the database, the researchers suggested they were encouraged by the ubiquity of the PDA, given its ability to make tools such as the DDI database available to

many people in many places. Research in Korea looked at PDA-based structured form-filling combined with a DSS to improve guideline adherence and better decision-making (Lee et al., 2009). The control group used a PDA-based form for clinical encounters and the experimental group used a PDA and DSS. The researchers' findings provided evidence that the use of the DSS both increased the likelihood that the correct obesity-related diagnosis would be made and the likelihood of a missed diagnoses was decreased.

Theme 5: Device used to access other health systems. One of the studies referred directly to EHR access via mobile devices. The study describes a web-based application to provide patients with access to their cardiology records via a portal to a commercial EHR system. Although interviews with participant patients revealed varying levels of comfort using the system, the patients were consistent in their enthusiasm and acceptance (Vawdrey et al., 2011). The paper stated that the study was done to provide patient access to their EHRs on a tablet. While the authors' claims may have been true for a tablet, since this was the only paper in this review that did so, it was reminiscent of work done by Cimino (2002) who described a similar scenario involving the use of a PDA-based EHR portal named PatCIS.

Two papers discussed medication management applications. The first was named the Colorado Care Tablet in Siek et al. (2010), an application designed for older adults to manage their medications. The study reinforced the notion that usability testing needed to be done within the context of use, as was described in Svanæs et al., (2008). The other study described an evaluation of a PDA-based application that provided prescribers with their patients' prescription histories in an effort to mitigate adverse DDIs. They gave 1615 prescribers access to the database of a 100-day patient-specific medication history

(Malone & Saverno, 2012). Unfortunately, despite the addition of e-prescribing and automatic DDI checking features, prescribers' use of the device waned over time and the study concluded that the use of this device did not affect the rate of adverse DDIs within the group. The paper did not mention any usability testing of the application before it was used in the study and the state of the app changed considerably during the study.

The remaining papers in this group referred to clinician-used EHR portals. One paper evaluated software for home-help service staff, and found only minor differences in input efficiency between novice and experienced users (Scandurra, Hagglund, Koch, & Lind, 2008). Another paper reported that tablet computers in an ambulatory care clinic were well-received by clinicians (Murphy, Wong, & Martin, 2009), and another advocated for simultaneously testing of many prototypes for use in a clinical environment (Karahoca, Bayraktar, Tatoglu, & Karahoca, 2010). Research from the United States investigated PDA-use during rounds by measuring the time saved compared to rounds accomplished without PDAs. Before and after, test results for a group of 22 residents showed that the task time, i.e., time to complete a task, dropped from 50 minutes to 40 minutes when PDAs were used (Park, Tymitz, Engel, & Welling, 2007). The study concluded that residents were better organized with their PDAs.

Three papers, two from Norway, and one from Canada, described a testing method that extended standard usability GUI investigations by including human factors and ergonomics. In Svanaes, Alsos and Dahl (2010) the researchers used a full-scale usability lab in the form of a bedside simulation, and tested a handheld device that could remotely control a patient touch-screen terminal. The mobile device acted as a controller for the terminal and allowed the clinician to discuss medical matters with the participant

while using the terminal as a visual aid to display imaging results. Similar to Karahoca et al. (2010), in this study many design prototypes were tested and alternatives compared. Participants noted that the terminal screen was large enough to clearly see the images, but that the mobile device was too small for that purpose. The combination of a PDA as controller and terminal as display, allowed for better doctor-patient communication and influenced the participants' perceptions. This resulted in participants favouring a PDA-display solution over a non-PDA solution. The second paper from Norway by Alsos (2008) examined usability issues within the context of attention theft, a concept whereby the use of a device demands so much of a clinician's attention, that it disturbs the communication between them and their patient. This may be quantified by measuring the user's focus shifts and episodes of slower speech during an encounter. In a study conducted in Toronto, a verbal protocol analysis of the usability of a mobile EMR asked: given that new technology requires behavioral change, did the PDA offer the portability of paper and the information on demand from the EMR? The researchers' answer was a succinct "not yet" (Wu, Orr, Chignell, & Straus, 2008). Of interest in this paper were the quoted examples of participants interviews. Regarding mobility: "I was able to get it and if I would be walking while doing it there would be no problem"; "The whole point is to save time". Regarding data entry and form factor: "That's it, give me a keyboard, no handheld for me"; "So off go my glasses, since I cannot see your device here"; "The patient died because you lost your little stylus there" all from (Wu et al., 2008).

Two papers covered order-entry via tablet (Dawson & Kushniruk, 2009) and a PDA (Zwarenstein, Dainty, Quan, Kiss, & Adhikari, 2007). In Dawson and Kushniruk (2009) six nurses entered Doctors' orders into a tablet computer during a usability analysis, and

generic and application-specific strategies were used to mitigate usability issues. In Zwarenstein (2007), a study protocol description (no results) of a 65-week trial in which prescribers used PDA devices to order medications for patients via a centralized database server. The server was programmed to ‘turn off’ access to prescribers on a randomized week-on week-off schedule. From the prescribers’ perspective, the PDA would function normally, but the CPOE app would simply not work during off-weeks, forcing them to return to a paper-based workflow. The outcome of this study was not found in the literature.

The Role of Error in the Safety of New Computing Devices in Healthcare

This part of the literature review focusses on the safety and potential error-inducing characteristics of mobile handheld devices in healthcare when used to order medications and treatments. The author also reviews human factors in computing papers and technology induced error papers. The intent of this literature review is to answer the following question: Is there a correlation between the size of the handheld device’s display and the perceived safety of the iPhone® and the iPad®?

Methods

Search strategy. A search was conducted for papers on the general use of technology and its problems in healthcare and in other areas. The PubMed database was searched using a set of terms for order entry and for the concept of technology errors to construct the following Boolean query:

(AND technology AND human AND factors)
AND (CPOE OR (Computerized AND provider AND order AND entry)))

AND (error Technology AND Induced AND Error)

Specific terms for the devices were not included to allow for a more general collection of results in the topic. The query was limited to English language papers published from 2007 to 2017. The year 2007 was chosen to align the results on the same timeline as the other sections of the literature review.

Review of identified studies. Papers that met the inclusion criteria outlined earlier were retrieved for full review. Papers were then examined for redundant findings and duplicate results. The set of full-text papers that remained was reviewed in full. Relevant themes and findings were extracted from the included papers and are grouped, presented, and discussed in the results section of this review.

Results

On review, three major themes emerged on errors in, safety of and interaction with computing technology: (1) how technology affected human performance, (2) trust in technology, and (3) other human factors. The author will discuss the themes in the next section.

Theme 1: how technology affects human performance. Responding to concerns expressed in the literature about applications being unsuitable for clinical use, Lilholt and colleagues (2006) devised and tested a usability evaluation to investigate problems in an EHR system by combining methods from high-quality laboratory simulations and field studies (Lilholt et al., 2006). The researchers concluded that some of the usability issues they encountered only manifested themselves in realistic settings. Their findings suggest that if the scenario were not realistic, some usability issues may not have been

encountered. Some simulations were limited by their lack of realism, insofar as the consequences of participants' errors were not immediately apparent. In a highly realistic simulation, with an actor playing the part of a patient, in a study that examined aspects of the doctor-patient encounter, Lilholt argued, this level of realism could affect the participant's performance. Authors reported high quality simulations i.e., (Santos, Teixeira, Ferraz, & Carvalho, 2008; van der Sijs, van Gelder, Vulto, Berg, & Aarts, 2010) led to valid results, had fewer stated limitations, were more effective and left little room for doubt or criticism of the approach to studying errors. Yee, et al., (2006), in a paper from Australia, described a holistic view of the impact of a new technology upon medical handovers and acknowledged that real-world studies were considerably more difficult to conduct, but and were necessary because the complexity of interaction between people and machines, i.e., human factors and usability were important (Yee et al., 2006).

In a non-medical paper, Sasson and colleagues (2006) used a simulation approach to combine a Human Performance System model with Applied Behaviour Analysis to measure the effects of a changed work process on human performance improvement. Alternatively, Lilholt and colleagues (2006) stated that the results of simulation studies may not be equivalent to the findings from studies conducted in a more naturalistic setting. In a non-medical paper from Brazil, dos Santos (2008) investigated the safety-conscious nuclear power industry. The researchers studied the design of control room instrumentation. The report concluded that in applying the principles of usability to the design of new interfaces, the operators spent less time identifying the type of simulated nuclear accident that was in progress, and that navigation among the new interfaces was

more efficient. The simulator used in this research duplicated the control room of a real nuclear reactor. This scenario was a relative luxury for usability researchers.

An airline industry study asked: “would the increasing reliance on automation in air traffic control affect the safety culture? With the greatly changed role of automation and the improved systems development techniques – errors that humans have learned to recognize may become rare; and error conditions that have never existed before, might now be encountered” (Zemrowski, 2008). A key concept from this paper was the inevitability of increasing automation; this, because studies have shown that an increase in air traffic levels would exceed humans’ ability to handle the additional activity, that relies on people to distinguish one aircraft from another, using radar and other modern surveillance. As automation would enable people to do more than is humanly possible, what would happen if the technology failed? Implicit in this statement is that one trusts the technology. The paper concluded that new frameworks and methodologies were needed to study automation.

An example of a health-related, in situ, in-depth study was described in Miller et al. (2010). For five days, five patients’ clinical bedside conversations were recorded and observed over time by an embedded researcher who investigated the merits of paper as a Computer Information System (CIS) technology prior to a Computerized Physician Order Entry (CPOE) installation. The researchers showed that different clinicians used paper in different ways. Although paper and pencil were the same for all clinicians, the methods of, and reasons for paper and pencil use varied amongst the participants.

Theme 2: trust in technology. An English poll of safety, quality improvement, and health-care administrative leaders identified health information technology safety as the

hazard of greatest concern (Denham et al., 2013). The researchers advise medical practitioners to trust technology. They urge prudence and recommend a strategy of “trust but verify” in their conclusion. Trust in technology has been an emergent theme in this review. A poll A short case study by Nolan (2008), told the story of a medication error in which a physician using a remote CPOE workstation, ordered the correct drug, but for the wrong patient. The medication order, a paralytic drug, was dispensed and administered with dramatic results, when the patient experienced the effects of the paralytic while walking to the lavatory. Having lost control of his limbs, the patient collapsed onto the floor. He was cared for and recovered shortly thereafter. The physician’s order was carried out, checks were made, all was in order except that no staff noticed or questioned the inappropriate medication for the patient. The physician’s decision was trusted, and since the treatment was ordered remotely, the trust was unquestioned by five staff members in the chain, all the way to the patient. The technology made no error, but was unable to detect the error. This particular CPOE system had not been set up to check for appropriateness for patient condition. As a direct result of this case, the ward changed its procedure and medication labeling.

In Montague, Kleiner and Iii (2009), the notion of trust in technology was investigated. “It is not apparent if trust in medical technology is the same as trust in technology. If the two constructs are different, existing trust in technology models may not be useful in discussions about medical technology (Montague et al., 2009).” Montague and colleagues’ paper (2009), with its elaborate set-theory diagrams, sought to understand how people constructed trust in medical technology compared with the same constructs in other general technology. The researchers showed that participants believed that 73% of

words for ‘trust’ on medical and general technology were the same. This, according to the paper’s authors, was not sufficient to say that participants trust general technology and medical technology in equal measure. The words that differentiate trust in general, and medical technology, are ‘error’ and ‘hesitation’.

Healthcare team members did seem to trust paper. Saleem et al. (2009) conducted a study to explore and understand why people relied on paper and paper-based workarounds, even in medical centers with fully implemented EHRs. The authors conducted 20 semi-structured key-informant interviews with a broad range of clinicians, administrators and IT specialists. Interviews were constructed using a guiding framework of predicted sources of paper-use. The study identified several reasons, for using paper. Some of these reasons were: task-specificity, personal preference, efficiency, trust, ease of use, memory, data-organization, awareness, task-complexity, and security. The authors concluded that paper, in these digital times, remains a valuable and useful technology. They also concluded that in some cases paper facilitated errors and gave an example: when staff made use of paper to work around problems they encountered when using an EHR. They note that not all the information recorded temporarily on paper was replicated within the EHR.

These four papers covered paper-use vs CPOE and its ability to aid nurses to detect medication administration errors (Sowan et al., 2006), paper’s role in medication administration process (Cunningham, Geller, & Clarke, 2008), the impact of paper on communication (Pirnejad, Niazkhani, van der Sijs, Berg, & Bal, 2009), and the ability of paper to affect overall medication error rates (Devine et al., 2010). All of the researchers

discussed aspects of paper-based ordering vs CPOE in implementations of the new technology.

Theme 3: other human factors. In 2009, six papers were published describing methods used to study technology and error. The proposed methods differ from the in-situ or high-realism methods described above. All but one paper was published by authors affiliated with the University of Victoria Department of Health Informatics (UVic), to deal specifically with research methods to test for (Borycki & Kushniruk, 2009), diagnose sources of (Borycki, Kushniruk, Keay, & Kuo, 2009a), use heuristics to (Carvalho, Borycki, & Kushniruk, 2009), predict (Borycki, Kushniruk, Keay, Nicoll, Anderson, et al., 2009b) and prevent (Kushniruk et al., 2009) technology-induced error in healthcare. All of the researchers advocated the utility of using mixed methods and simulations in this research, primarily to answer the question ‘why’ in order to establish root causes of technology-induced error. The authors found that simulation methods could be used with success in evaluating clinical hardware and software. Using heuristic evaluations, researchers may predict the likely error-rates of systems before they are procured and deployed in hospitals. These tasks may be accomplished using the provided frameworks and methodologies, developed and refined over many years, at minimal cost to the research team or organization undertaking such studies. The sixth paper, from the Veterans Affairs in the United States, contributed to research methods introducing a qualitative human-factors protocol to evaluate systems and processes thought to lead to diagnostic and medication error (Hysong et al., 2009). This new protocol described

methods derived from the psychological sciences and human-factors research to analyze the ways CPOE users communicated and identified problems.

A theme, implicit in all the research-method studies was the need for more usability studies to be done on the ever-growing list of available electronic medical devices, hardware and software. “Future research should include usability studies to investigate how alerts should be presented to be safe and acceptable to clinicians; i.e., several alerts in one pop-up, clear and concise alert texts, nonintrusive alerts” (van der Sijs, Aarts, van Gelder, Berg, & Vulto, 2008). “More interdisciplinary work is needed to ensure that clinical systems are designed for maximum benefit of all stakeholders, to increase understanding of information needs and requirements across settings, and to understand shared user performance with devices” (Alexander & Staggers, 2009). “The usability of medical devices has become a main issue for patient safety” (Röllerke, 2009). “We evaluated a CPOE system after approximately three years of its successful implementation and use, and found many interoperability problems that led its users to adopt error-prone compensatory strategies” (Pirnejad et al., 2009). According to the reviewed papers, many clinical tools, devices and systems were not tailored to users’ variability. User interfaces were criticized for their inflexible login modes, medication models, controls, (Wenzer, Böttger, & Boye, 2006); inflexible processes (Miller et al., 2010); ordering formats (Jayawardena et al., 2007); rigid designs (Khajouei, de Jongh, & Jaspers, 2009) and rigid workflows (Borycki, Kushniruk, Keay, & Kuo, 2009a).

Devices, software, machines and interfaces were only part of the usability study. Human factors have been discussed in several studies. Results from human-factors studies have demonstrated that different people use the same tools in different ways. The

authors suggested that the differences in users' life experience, risk-taking, and cognitive processes have played a part in how decision-support tools are used (Miller et al., 2010), that is echoed in the findings of Goodman and Miller (2006). Miller reported that nurses and doctors used paper tools, i.e., charts, in different ways due to their different roles on a ward, indicating the need for role-based flexibility in CPOE design (Miller et al., 2010). Arif studied the differences between novices and experts' use of a user interface and, noted inter alia, differences in error rates and recovery (Arif & Stuerzlinger, 2010). In Mahlmeister (2010) publication, novice and expert differences in task completion were also examined. People varied in their ability, skill, familiarity, competence, confidence and style.

Discussion

Some errors cause new learning experiences that can lead to scientific progress. Other errors are more of an annoyance and, if made in a healthcare delivery context, may have harmful consequences. Mobile devices, a new technology, play a part in facilitating harmful errors (Kushniruk et al., 2009) at the same time as they improve efficiency in the workplace (Park et al., 2007). Since human error abounds, it must be managed. To manage an error, it must first be measured in context. This can be achieved through usability evaluations and analyzing specific metrics and outcomes.

Mobile devices proliferate. Mobile devices and PDAs are ubiquitous and much effort has been made advocating for their continued use, testing and acceptance. Much of the research discovered in this review reports positive results, improved worker efficiency and general enthusiasm. These machines are here to stay and reports of poor performance or troubles are not met with warnings against their use. Instead, clinicians

are met with directions for improvement and at worst, a delay in applicability pending the next generation of devices. Even the naysayers are optimistic and are motivated to re-engage with technology once the limitations are no longer relevant. This enthusiasm may be shored up by the relatively short development cycle. Consumers are experiencing high technology marketplace. In fact, the data-collection aspect of this research, i.e., the methods of screen capture, was directly affected by changes in technology. What was formerly a difficult task is currently greatly simplified allowing for a much simpler and higher-quality data-collection method. There is a notion that if a technological task is difficult to achieve presently, then in time improvement to the technology will lessen the difficulty and make our efforts more productive. This is not a new idea. Personal computing has been in a process of continuous improvement ever since the introduction of the IBM-PC in the 1970s (Wikipedia, n.d.).

The choice of a smaller or larger display is not exclusively a matter of aesthetics and appealing designs. There may be good reasons for keeping displays small, as for example in the medical context, in which the devices must be small to meet intimacy and/or acceptance demands (i.e., signal warning devices for blood pressure or chemistry). (Ziefle, 2010)

The form factor of a small mobile device affords some body language that can contribute to a positive doctor-patient encounter.

One prediction that may be made with some certainty is that the use of mobile devices in healthcare is set to grow. Until such time as the capabilities, user interfaces and form factors stabilize, then the research that examines, tests and makes determinations on a specific device then that research is only really valid for that device and cannot be generalized to other devices, even if the other device uses the same software. The mobile

device is not in and of itself a useful tool unless the design and usability of software running on the device and the capabilities of its user and the underlying infrastructure is also considered. A systems-thinking approach is needed to assess mobile device usefulness.

Usability assessment of mobile devices is difficult due to frequent updates. It is difficult to generalize the usefulness or applicability of a mobile device found in this search of the literature. Many of the problems encountered in earlier papers were due to limited capabilities of older devices. Newer devices solve some older problems but also introduce new issues. For example, researchers who were frustrated by the requirement that the Palm device needed to be physically connected to a server to download its data, were troubled when devices that contained important health data were lost before data synchronization could be done (Gallagher et al., 2008). This problem can be solved by using newer devices, that are always connected to the network, allowing researchers to collect data in real time. Another example is found in Haller (2009), in which the researcher urges caution when using handhelds for patient questionnaires due to increased data-entry times because the device required the user to use an adapted lettering scheme, called Graffiti. Graffiti is error-prone until the user develops the skill to use it efficiently. The PDA device was rendered obsolete and has been superseded by a new class of devices that use a multi-touch UI rather than a stylus-based UI. In this case, the problem of graffiti errors is replaced with errors in the navigation of a multi-touch UI having to develop skills in entering text on a virtual keyboard. The same principle also applies to the usability tests that report positive outcomes. The usability study results

may not be fully useful when applied to newer mobile devices that have radically different data-input modalities.

System-level software and apps are also frequently updated. Given that the text input modality in most current mobile devices is software-based, modes of input can occur altering the GUI in subtle ways, i.e., the behaviour of predictive text entry, or in dramatic ways, such as a new split-keyboard layout for larger devices that require the use of two thumbs rather than all fingers to enter text efficiently. Another example is a change to the iPad® system software, in iOS 5, where a four-finger gesture would cause the currently running application to be moved off the screen and be replaced by the previously running application. This gesture is a potentially useful feature but is easily made in error. Unless the user is both familiar with it and how to recover from an error, the result could lead to confusion, distraction, and errors. These usability changes can be caused by a routine software update. Many devices are set by default to have software updates occur automatically. These examples show that a ‘minor’ change, i.e., same application, same device different OS can invalidate a usability score and cause problems or the need for further training. As a usability test result is only valid for a specific device state and because the device state can change frequently, then the only way to keep up to date would be to test for usability often. Testing rapidly may facilitate frequent testing. Testing frequently may be feasible if the testing were simple and standardized.

Usability research studies, despite the high likelihood of obsolescence at the time of publication, remain necessary, perhaps not for the specific result’s future usefulness, but more for the utility of the methods used to achieve those results, particularly the metrics.

Implicit in all the research-method studies in the selected papers, is the need, expressed with urgency in some, for more usability studies to be done on the ever-growing list of available electronic medical devices, hardware and software. All technology-induced error studies reviewed here advocate for the utility of using mixed-methods primarily to answer the question ‘why’ to establish the root cause of errors. Quantitative studies have established the fact that errors occur and occur at a greater rate when procedures change. Qualitative papers (n=17) were fewer in number than quantitative papers (n=24). Mixed method (n=10) studies employ qualitative and quantitative approaches and are popular to use in the study of technology-induced errors.

“[...] that devices which are already on the market differ drastically in their usability [...] and] that unsatisfying usability bears a high risk in health care.” (Büchel, Baumann, & Matern, 2007).

As the research-methods papers show, the study of technology-induced errors need not be an expensive or a labour-intensive endeavor. Some expertise is needed to perform such tests, but suitable trained IT staff using resources described in this review may also quickly learn this and further, learn how to standardize testing for a particular health organization.

It is interesting to note that the U.S. Food and Drug Administration (FDA) now requires usability testing for medical devices (Alexander & Stagers, 2009). The FDA also mandated that device-related deaths and serious events be reported to itself and to the manufacturer (Mahlmeister, 2010). There is discussion of mobile devices’ software being classed as medical devices by the FDA in Dolan (2009). If this becomes policy, software would require agency approval before its legal use in healthcare facilities. Agency approval would depend on a successful usability assessment. With the growing

demand for mobile devices, the expected requirement in the United States by the FDA for usability testing, the clear need for quality assessments of devices and device-software pairs, and the knowledge that untested, use of these devices can lead to new errors and harm, has made the need for usability assessments more important. Research must be done in this area for the safety of healthcare everywhere these devices are being used. The scope of this usability-testing problem is larger than it may seem given the current proliferation and global distribution of new, sophisticated mobile devices. The recommendation here is for usability testing methods and is a pragmatic approach.

That one device proves to be usable, does not mean that another similar device, or its software is automatically similarly usable. Ideally, every device-software combination should be tested for its error-inducing potential as part of a procurement process. As the research-method papers show, this need not be an expensive or labour-intensive endeavor. Expertise is needed to perform the tests and suitably trained IT staff may quickly learn this skill.

Mobile device usability is context-sensitive. Usability, acceptability, usefulness, and cost-effectiveness are all concepts sensitive to context. For example, a usability assessment for a tablet-based patient background assessment, to be completed by the patient while they were waiting for an appointment to begin, was usable for most patients. The researchers noted that a small subset of their participants, a separate cohort has quite different usability results (Hess et al., 2008). When tested in the general population, the usability of these devices was good. When tested in a different context, a cohort of older individuals, usability was poor. If this tool is intended for both environments, then two assessments needed to be completed, one for each context.

Another example is a mobile image-viewing application assessed for usability in an imaging suite may not be as usable when used in a brightly lit emergency room.

There is no standardized usability or assessment method that is suitable to test and evaluate all cases because each case is unique and requires unique solutions. This concept is echoed in the papers examining safety that show different people use data recording tools, such as the pencil and paper, a smartphone or a PDA in different ways. Given that these devices, both old and new, can be seen as tools in a kit of useful items, a similar approach should be taken for testing the utility and usability of these tools. A knowledge-based kit consisting of frameworks, heuristics, metrics, techniques and analytical methods is needed to be able to create an effective evaluation of a given system. It is imperative that the resulting evaluation be done in a timely manner because the result will only be valid for a short time due to the rapid and frequent update cycles for a mobile device's software and hardware.

Rapid changes and multiple contexts require agile test methods. The mobile device software development life cycle is rapid. The operating system on the iPhone® and iPad®, *iOS*, had major releases each year for the past six years, each with significant new features that alter the way the device is used (“iOS version history,” n.d.). Many will be subtle incremental updates, but some will change the way the device is used. Such significant changes invalidate any usability investigations based on previous versions of *iOS*. The pace of software development and major new items may or may not slow as the technology plateaus. The competition for market share in the mobile device arena is

likely to remain fierce prompting the introduction of new devices and new methods of interaction.

Realistic simulations provide good insight, but they also are slow to build and studies that are not rapid and may not be able to keep pace with major software changes result in more papers like Galliher's (2008). Galliher (2008) found that devices' design limitations affect their utility. The researcher suggested that future devices would likely be much more usable. Since then, any study reporting device-specific findings may potentially be dismissed with a statement such as "Well, that was then. Now, with the new version of X things are different", in other words, if one does not like the state of their phone's software, one can simply wait a few months and it will likely be different. Slow methods of usability and error-testing are themselves less useful now than they were in the past. What practical use is an error-result that was measured by a device that has changed fundamentally, due to a major improvement, invalidating that result on publication? The answer, perhaps may be the methods and metrics.

Fast methods that enable a rapid turnaround are essential to have results meaningful to both the community of users, our clinicians, and to the product's developers. In other words, if research is to be seen by the developers so that it is relevant and useful to improve the usefulness, it needs to be conducted rapidly and in a timely manner so that it can be used to improve the product before the product reaches its next major release. Slow and fast are terms that are relative to the pace of change of technology. For example, a mobile EHR app running on an iPhone® that a researcher investigates for its error-inducing properties will have a short window of opportunity to publish, and have

the published data remain valid before the date of the next major release of either the software or the device's operating system.

In healthcare, mobile device adoption is increasing (Raptis, Tselios, Kjeldskov, & Skov, 2013). The devices' software is improving regularly and the number of available software apps is increasing. These apps are appearing on devices of different types. Different types of devices have different modes of use, i.e., some use a touchscreen software keyboard, some a hardware keyboard and others use a stylus. Now, some can perform accurate voice dictation. This march forward results in a set of device-software items that seem to be growing exponentially for the sake of a hypothetical example, say we have 10 devices, 25 000 apps related to healthcare², a software release cycle of 2 per year and a major device system software release once annually ("iOS version history," n.d.). The product of numbers reveals that there are half a million device-software combinations. Given the current frequency of major releases, any tests done on these pairs would likely be invalidated two to three times a year.

The Problems with Smartphone and Tablet Form Factors

The previous parts of this literature review focussed on the safety and error-inducing characteristics of mobile handheld devices and the usability and usefulness of these devices and their impact on safety and error-inducing characteristics in healthcare. This final part will review papers about the general problems and difficulties that users of

² An approximate total number of Medical and health & Fitness apps from Apple's *App Store* retrieved 2016-02-28

these devices encounter. The intent is to discover more about the devices' usability not necessarily within the context of healthcare.

Search strategy. A search was conducted for papers about the effect of screen size on the usability of mobile devices. The *Google Scholar* database was searched using a simple query:

mobile device AND usability AND ("effect of screen size" OR "effect of display size")

The query was limited to English language papers published from 2007 to 2017. The year 2007 was chosen to be consistent with the previous two searches. Papers were not limited to healthcare.

Review of identified studies. Papers that met the inclusion criteria outlined earlier were retrieved for full review. Papers were then examined for redundant findings and duplicate results. The set of full-text papers that remained were reviewed in full. Relevant themes and findings were extracted from the included papers. These are grouped, presented and discussed in the results section of this review.

Results

These papers all report on studies of users interacting with devices. They may be categorized into four themes: (1) Passive mode, (2) Active mode, (3) Large Displays, and (4) Multimodal activity. The remainder of this review will detail the findings from the simplest to the most complex study of user interaction.

Theme 1: passive mode usage. Passive mode here means that the user of the device was engaged in simple navigation and reading or watching content on the device rather than engaging in active interaction. In a paper by Maniar (2007), the researchers

examined the effect of learning on the efficacy of students' test performance.

Specifically, looking at the effect of screen size on: learning performance, the effect of the delivery media and the students' subjective performance. The subjects watched video and read pre-formatted documents on a cell-phone display and the other group did the same on a PDA display. Subjects were interviewed for their opinion of the process. The results show that both media were more effective on larger screens, the text-based content was more effective than video content on the students' performance. Despite the improved performance, there was also evidence of a preference to watch video rather than read on all devices even though it was less effective on their scores (Maniar, 2007).

Using larger screens, Raptis et al. (2013) conducted an investigation into the usability and effectiveness of mobile devices as affected by screen size. The researchers found that not all tasks benefitted from a larger display, particularly those that did not require a lot of scrolling, (Raptis et al., 2013). The screen size of mobile devices did not have a significant effect on usability scores; however, the participants' prior experience and desire for a specific device did have a significant effect. The study suggests that there is a minimum size of about 11cm for the display at which the benefit of a larger screen on effectiveness and efficiency improves. This is consistent with the paper by Maniar (2007) where the authors found that efficiency and effectiveness results improve with sizes larger than 11 cm. The authors suggested that researchers measuring perceived usability using System Usability Scale (SUS), a tool that provides a set of heuristics for measuring usability as explained in Brooke (2013), may not observe significant differences when evaluating the same app on different screen sizes.

In a paper from Denmark (Kanstrup & Stage, 2009), researchers describe a case study on the usability and utility of training on a PDA-sized device. The researchers were interested in using the PDA to replace the average of about two-and-a-half kilos of reference material junior doctors carried in their lab-coat pockets. They asked the question: “Could the PDA replace some or most of this bulk?” After building a functional prototype and investigating its use on the ward, the authors concluded that the system was of limited utility. This was due to the software being incomplete and users’ experiencing a high number of usability impediments, particularly in the number of steps users had to take to make progress in the navigation of the app. The authors call for further case-studies and no mention was made of experimenting with larger screen-sizes.

Two of the studies in this group made bio-medical measurements: (1) An eye-tracking device was used to measure the usability of a variety of common apps in mobile devices, (Al-Showarah, AL-Jawad, & Sellahewa, 2014). Results suggest that across all age groups, the device with the smaller screen was less usable than the device with the larger screen. In the other study (2), the researchers used a brainwave detector (EEG) (Chen & Lin, 2014). The EEG was used to estimate the users’ attention level while reading text on a small-screen mobile device within a number of contexts, i.e., sitting, standing, walking and so on. Different display methods, manual scrolling, auto-scrolling, paging were used within each context. This detailed study revealed that device-use under different contexts would require choosing the appropriate display method for that context for optimal results. The authors concluded that sitting while using a mobile device generated the highest sustained attention.

Theme 2: active mode use. Active-mode use refers to the users' interaction with the mobile device beyond simple navigation. For example, in a paper by Shrestha (2007), users' web-browsing performance was observed and measured on a desktop computer and a mobile browser on a cell-phone. Incidents such as: getting lost, complaining and cursing out loud were observed and recorded in a controlled setting. Participants were given a set of tasks and asked to think-aloud. Performance was poor. Participants used 80% of their allotted time to navigate the mobile browser and only 20% of their time to do the same on the desktop. This result is unsurprising and perhaps not that relevant to current mobile device form factor experiments because this study's publication predates the implementation of responsive-design technology in the target website, i.e., websites can now behave very differently when displayed on mobile browsers, and because the current set of touch-based user interfaces on mobile devices are significantly different from the ones used in the study.

Research from Indonesia investigated the effect of screen size on usability by making and using an app designed to test ergonomics (Restyandito, 2017). After using the app on four different screen sizes, they found that screen size and the user's hand size affected error rates and usability. Hand size had the most significant effect when thumbs were used to make gestures on the screen (Restyandito, 2017).

Towards a new design, researchers at University of British Columbia sought to measure the effect of an adaptive GUI scheme, where recently used items are shown first in drop-down menus on the user's performance (Findlater & McGrenere, 2008). A complex study with many variables, achieved its goal in part by making the accuracy of the adaptation a variable so that at times the adaptive GUI was more usefully accurate'

than at other times was deliberately flawed. A desktop-sized computer was used as a reference. The results showed that the users' performance on a mobile browser was slower and that the effect of the adaptive accuracy on the users' performance was greater on the small screen.

In a study led by Valdez (2010), the authors, interested in producing a diabetes self-care application for home use, undertook a small-scale study that measured the learnability differences between a large back-projected wall touch-screen and a small handheld-sized GUI presented centered on a 15" laptop screen with a plain border. The GUI was the same, just displayed at different sizes. Each participant used each device. Half of participants started on the device with the smaller display. The other half started on the device with the larger display. Researchers concluded that users of the two very different sizes of display were equally effective in getting their tasks completed. All users' performance improved, i.e., required less time and effort, when using the second device. Performance improvement was greatest when the large display was used first. Users gained an unclear benefit by using the large screen prior to repeating a similar task on the small screen. The qualitative portion of the study reports that the users' interaction with the large screen was perceived to be easier than the smaller screen (Valdez et al., 2010).

Results are presented in a paper by Dey (2012), on a series of experiments on depth perception in augmented reality using an iPhone® and an iPad®. These tests make use of the devices' ability to know of their location and of their motion relative to their current position. Depth perception was either egocentric, the distance from the viewer to an object, or exocentric, the distance between two objects. By overlaying a graphic object

onto the video image from the devices the researchers evaluated the participants' depth perception. Results showed that egocentric depth perception was equivalent for the two devices and that there was a strong user preference for the larger screen. The exocentric test, they found a significantly lower error using the larger iPad®.

Theme 3: active mode use on larger displays. Tseng, Chao, Feng and Hwang (2013) studied display size against search performance during map reading. The displays were 7", 15" and 21" square on which was displayed a military map with battlefield icons. Participants were engaged with a series of tests involving visually searching the field of view for icons. Their results showed that larger icons resulted in shorter visual search times and that smaller screen sizes had longer search times. The longest search time was that of the largest screen combined with the smallest icons. The participants' accuracy increased with the larger screen sizes. The authors concluded that the intended use of the device should be considered when deciding which device to use. For example, if search speed is the main concern, say in the field, then a small screen may be the best choice. If greater accuracy is desired, then the larger display, say, in a control room, would be a better choice. This reinforces the finding that different people use the same tool in different ways as seen in Miller (2010)

Jakobsen and Hornbaek (2013) examined user's performance on: a very large 6-foot, a normal 20-inch and a small panel monitor while reading and interpreting geographical maps. In their two experiments, both indicated that the larger displays led to increased task completion times for their visualization techniques likely due to the time needed for the user's eyes to traverse the screen. In Bi, Bae, and Balakrishnan (2009) an unusually large wall-sized display (about 5m x 2m) was set up for general office-based computer

tasks. This was not a comparison with smaller display sizes as such, but had some interesting findings: (1) the users' preference favoured the larger display. (2) using such a large display resulted on distinct user behaviour patterns when managing windows within the display. Unfortunately, there was no data for comparing the efficiency or effectiveness of this system to a system with a standard desktop display. As seen in previous studies, for example (Maniar, 2007) , users preferences did not always align with effectiveness (Dey et al., 2012).

Theme 4: multimodal activity. In a paper from the United States, researchers used the iPhone® and iPad® spatial orientation feature, i.e., the device senses its own position and spatial orientation, to be a gestural user input system (Spindler, Schuessler, Martsch, & Dachselt, 2014). Rather than having the user touch the screen, the device was controlled by moving about a controlled setting, an empty room, letting its location awareness control the GUI. The researcher received positive feedback from participants and learned that the larger screen had a faster completion time likely due to the reduced search times since the larger area displayed more of the user interface at any given time. The authors concluded with the notion that spatial input could be a useful augmentation to traditional finger-based manipulation. They also explored the benefits of haptic feedback.

The U.S. Army research laboratory published a paper by a team led by Redden, (2008) in which three displays: a small laptop, a tablet, a PDA and one goggle-mounted virtual display setup to be perceived as a desktop-sized display were used to remotely control and manoeuvre a military robot around an obstacle course (Redden, Pettitt, Carstens, & Elliott, 2008). The robot was equipped with cameras that the operator remotely viewed the environment using the display. The users' performance was graded (i.e., their

qualitative responses were collected on a Likert-scaled questionnaire.) The objective was to use the selected displays to operate the roving devices because the small size had a portability advantage over the robot's usual bulky controller. The task had greater complexity and a greater cognitive load than simple web-browsing challenges. The researchers measured how much effort the participants made by using a subjective workload assessment tool called the NASA TLX. Over 35 tables of data were presented and very few showed statistically significant differences between displays in the measured performance of the participants. The user preference was for the smaller standard PDA and tablet devices. The least preferred was the goggle-mounted virtual display where the users commented that the device caused eyestrain, the monocular vision was awkward, and they kept trying to move their heads to see a different view. Interestingly, it also had the highest measured workload index and required the most effort to use. Overall the participants preferred the smaller devices, the PDA and tablet sized devices and preferred the goggle-mounted device least. Unfortunately, unlike Maniar, (2007) where the preferred device was not the best performer, there is no meaningful data to show which device size elicited the best performance, only that the goggle-mount was the poorest performer. The overall finding was that the 3½ inch display (PDA) was as effective as a larger display for driving a small robot. The question remains whether the devices were sufficient for navigation and control.

That question was investigated in a follow-up study, by the same U.S. Army research laboratory team, where navigation information was shown to the user via a split-screen on a tablet-sized device or via a single PDA-sized device that toggled between two views or a third combination that used a vibrating belt as a navigation aid (in this case when the

system detected an obstacle on the vehicle's left side the belt vibrated on the operator's left hip) (Redden, Pettitt, Carstens, & Elliott, 2009b). The results of the multimodal sensory apparatus was published in another paper later that year focusing on the practical benefits of using more than one sensory mode by which the system concurrently delivers useful information to its user (Redden, Elliott, Pettitt, & Carstens, 2009a).

Discussion

Though not explicitly stated in Findlater's (2008) conclusion, the effect of an adaptive menu system appears to be more useful on cellphone displays due to the limited usable display space (Findlater & McGrenere, 2008). Raptis (2013) found evidence that there is a minimum size of about 11cm for the display at which the benefit on effectiveness and efficiency improves. If a team is creating software for use on a very small device, i.e., a cellphone less than 11cm, then including adaptive menus in its GUI could improve the usability of the system (Raptis et al., 2013).

Two of the studies in this group made bio-medical measurements, i.e., an eye tracking device (Al-Showarah et al., 2014), and an EEG (Chen & Lin, 2014). Findings from these suggest that clinicians who wish to be most effective, should avoid issuing orders on a mobile device while running on a treadmill and would be better to do so while seated. Interesting as this evidence is, these novel approaches seem not to improve the quality of measurement because the bio-medical methods offer only crude data, "the subject's brain is more active now than it was previously", while the more modest studies make measurements of cognition: "The subject understood and retained the information". Perhaps future technological improvements could present more meaningful results. In the meantime, the simpler research methods seem to yield more salient results

In Kanstrup and Stage (2009), we have another study on the usability and utility of training junior members of staff using a PDA. The authors concluded that the system was of limited utility. This study seems to contribute little more than to underscore the hidden complexity and magnitude of effort needed to create effective UIs. This is unfortunate as the goals were highly pragmatic.

The laptop used to simulate a handheld device in Valdez (2010) is an impediment to the goals of the study as they have very different modes of use, i.e., the handheld device has a touch-based interface that cannot be accurately mimicked on a laptop. The authors acknowledge these points in their discussions.

Spindler's (2014) study on corporeal movement as user-input is interesting, but there are some practical issues to consider: a user using the device in transit on a moving car may have the vehicle's movement trigger unwanted input or, a user in a crowded, confined or dangerous space such as a rush-hour train or an surgical operating room where large arm gestures could be impossible or undesirable. The notion of haptic feedback shows greater scope for investigation in medical practice. The U.S. Army's participants were given a highly complex task and rather than present all possible kinds of information visually. Participants were presented some of it in a tactile manner engaging another sense modality (i.e., touch) to provide data (i.e., collision detection) augmenting the visual display (Spindler et al., 2014). Could haptic feedback be useful for physicians to be warned of small errors (say, spelling mistakes) with a tactile modality? Would this ease alert fatigue or would it exacerbate alert fatigue?

Conclusion

Papers retrieved in this search reveal a wide range of experiments to both improve usefulness and to explore novel technology developments. Hospitals generate large amounts of data and healthcare IT departments can store and can arrange for its display. It is plausible there data can be displayed on large status boards. Radiological imaging can be displayed on many different screen sizes from a handheld phone to a wall-sized multimonitor panel display. The problems and novel solutions found in this review have a direct bearing on hospital data, monitoring and reporting. Given the unique setting of healthcare, i.e., controlled lighting and monitor calibration for medical imaging environments, the answers to the deceptively simple question: “Is a large screen better than a small screen in the emergency room?” may not be straightforward.

According to the papers reviewed here, the research is mixed and the findings are not generalized. Specific use cases, for example, a wall-sized electronic status board and an iPhone® view of it during a Code Orange (mass casualties) would require specific recommendations and further study. For example, if visual search speed is the main concern then a small screen should be better. If visual accuracy is preferred, then the larger display might be more suitable. If the user is processing a large volume of live data, then they may benefit from another sensory mode.

Chapter 3: Research Questions

The choice of mobile handheld devices available is constantly changing because the life-cycle of a mobile device is about a year between major product releases. Examining a particular product for its error-inducing performance would be of limited usefulness because, due to the lifecycle, manufacturers would yield a new device or major software update just in time for the results of a study on the previous version to be published in the literature. Greater insight may come from examining a more general property shared across all handheld mobile devices, a facet common to all devices that could transcend the product development life-cycle so that even if an older device it replaced the results could still retain some value. Properties common to all devices are things such as weight, input modality, i.e., stylus, keyboard, and screen-size. This research is focused on the latter.

Is There a Correlation Between the Size of the Handheld Device's Display and the Perceived Safety of the iPhone® and the iPad®?

The current marketplace offers mobile devices in many sizes. Mobile device sizes may be generalized into two groups: those that fit into one's pocket and those that do not. For example, the iPhone® easily slips into a standard labcoat pocket, but the iPad® is too large. One can imagine that, given the size difference, the methods of using them differ even if the user interface (UI) and the user input modality are the same, for example, the iPhone® can be used single-handedly for data entry even while the user is standing. The iPad®, on the other hand cannot be used easily with one hand unless the device is supported by some other means. Given these two general size classes, designers use the larger screens to optimize the UIs that take advantage of the extra surface area. Refer to

Figure 1, below, for an example of the relative display size difference of the user interface and the split-keyboard feature. With these variable physical properties, one can imagine that mistakes and errors made during use may also vary as will the user's preference for one or the other. Given the increasing complexities of user interfaces to healthcare applications, the question as to whether the form factor is important can be raised for such applications.

What Are the Preferences for iPhone® and the iPad® Form Factor in Clinical Use?

Different doctors use the same tools in different ways, i.e., a paper form is used to gather data in one scenario and the same filled-in form may be used in a different way for a group discussion in another scenario later the same day. Clinicians interact with the devices in different ways depending on both the context-of-use and the user's preferences. As the form factor differences are significant between these two devices, this may lead to a pattern of use for a particular clinician during a particular task for example, the iPad® may be preferred during group discussions or patient encounters because of its greater surface area and visibility. Other use-preference questions include: (1) are the devices used while standing? (2) are they used while seated at a desk? (3) and do users prefer to use paper as an intermediate step in a process? could be answered in an interview. Variable user interface design, form factors, modes-of-use and personal preferences can all contribute to the level of difficulties, manifested as usability problems, and to potential errors. By interviewing clinicians, we can learn what unique issues and problems are encountered by the use of mobile devices in a simulated clinical setting.

What Types of Usability Problems do Users have with Smartphone and Tablet Interfaces to an Electronic Medical Record?

The app is used by clinicians to access EHR data wherever they may be located. This data is usually reviewed from a standard desktop computer display. The novel appearance and interaction with the same data shown on smaller mobile devices will likely have new kinds of usability problems particularly with navigation as the modes of accessing information will differ from the standard workstation due to differing screen sizes and the touch-based interface. This final question will attempt to gain insight into the nature of new usability problems with the iPhone® and iPad® versions of the app. Knowing the answer to this question could influence the decision of which size of device to procure for use in a healthcare setting.

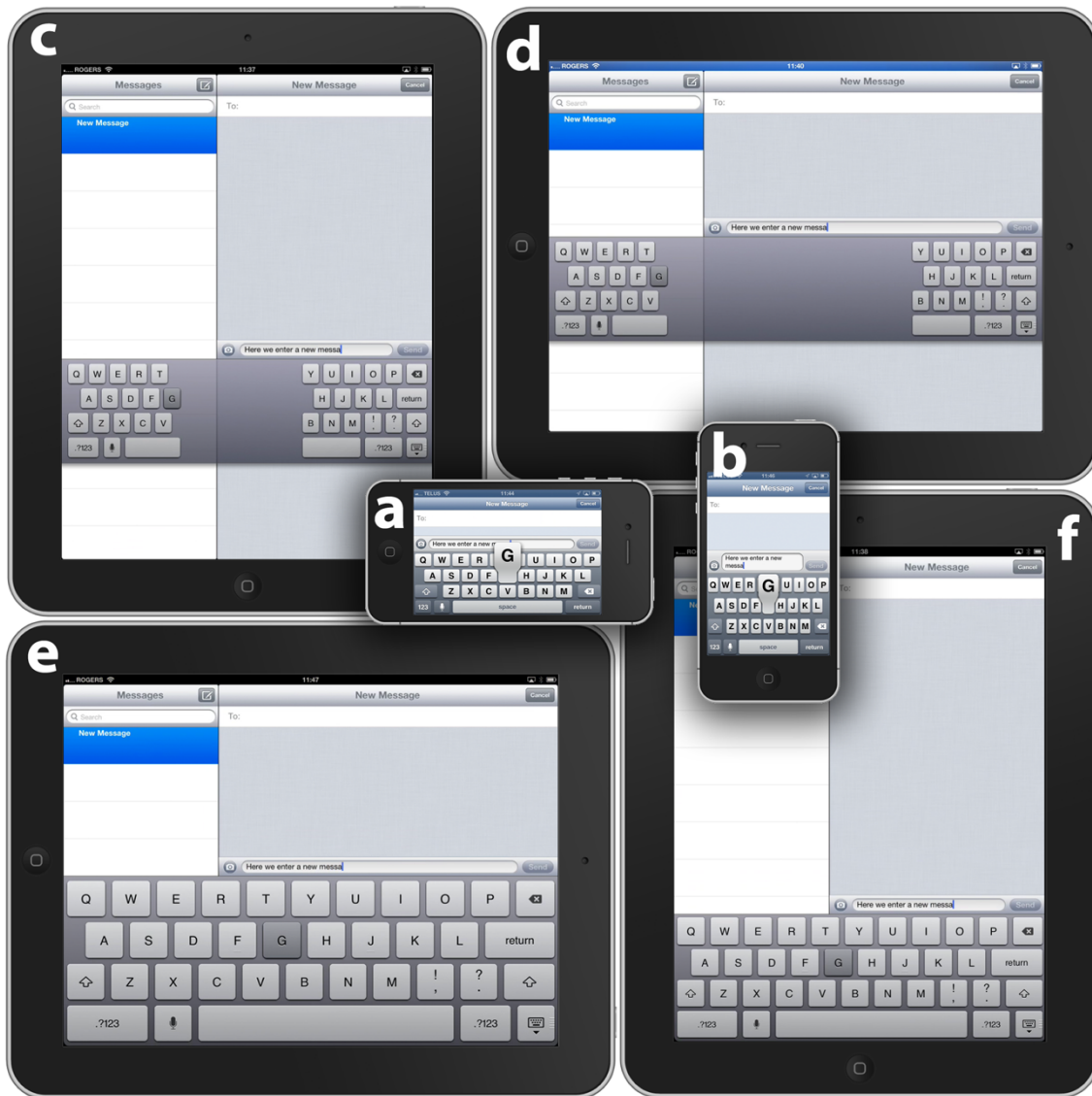


Figure 1: A side-by-side scale comparison of the iPad® and the iPhone GUI. Note that the letter G is being pressed on the keyboard area. The larger display affords designers more room to place visual elements. Notice the relative sizes of the keyboard input area and the different indicators for the currently pressed letter.

Chapter 4: Methods

This research extends previously published research by the University of Victoria in which a handheld device (PDA) and CPOE software were used in a simulation study by clinicians to enter prescription data for a range of medications (Kushniruk, Triola, Borycki, Stein, & Kannry, 2005a). In the Kushniruk et al (2005b) study, participants were in a room with recording equipment and a researcher working within the context of a realistic scenario. Screen recording was used to capture the users' interactions. Audio recordings were used to capture their think-aloud monologues. The recordings were analyzed to identify usability errors and actual medical errors made by the participants during the simulation. Usability issues were identified by analyzing actions that caused or were caused by problems users had during their interaction with the device. Medical errors were defined as errors of medicine, i.e., wrong medication, dose, frequency or other parameters made by the participant and potentially exacerbated by usability problems such as an inappropriate default value or by relevant information necessary to support physician decision-making being hidden off-screen. This published work identifies a link between specific usability problems and actual medical errors (2005b).

This research in a previous study examines error-inducing usability issues on the current generation of handheld devices, i.e., the iPhone® and iPad®. Compared with the PDA, the new devices had a different user input modality, display sizes and software. New consumer hardware devices in the style of the iPhone® and iPad® made by other manufacturers were not part of this research. The likelihood that they would also appear in clinical workplaces either as a medical tool, personal devices, or both, is high, and,

because they share a similar form factor, this research may also be relevant to similar device-software combinations.

A mixed-method design was used. Screen-recordings data were gathered during the participant's work session. Then, semi-structured interviews were conducted immediately following the work tasks, and again before the end of each session. The 90-minute session was conducted in a quiet room where the participant was allowed to work uninterrupted. The researcher was present to aid the participant in performing the task to prompt them through the stages of the research and to monitor the screen recordings as needed. See Figure 2 for an illustration of the audiovisual apparatus. For a more detailed explanation of the capture methods please refer to Appendix 7 on page 218



Figure 2: The research session screen-recording configuration.

The researcher's laptop was configured in such a way as to show the screen of the active device next to a video feed of the participant using the device. The camera attached to the laptop with a cable and the screen images of the mobile devices arrived over the wireless network. The whole laptop screen is captured to a media file on the external drive using screen-recording software

To gather evidence about the first question (i.e., Was there a correlation between the size of the handheld device's display and the perceived safety of the iPhone® and the iPad®?) Likert-based tools and open-ended interviews were conducted with the participant immediately following a session using each device and the app. The participant users performed the same task with each device, using a different case, each time and completed the questionnaire and open-ended interviews. When both devices were used and both interviews were completed, the participant was given a third questionnaire and interview. This final step allowed data to be gathered data on the second research question (i.e., What were the preferences for iPhone® and the iPad® form-factor in clinical use?) While each participant performed the tasks on each device, their actions and think-aloud monologue were recorded for analysis to gather evidence in order to answer the third question, (i.e., What types of usability problems did users have with the mobile device's interface to an electronic medical record?) Participant errors and problems were captured and analyzed in the screen recordings, interviews and Likert tools. This was of value because previous studies have shown that users' perceptions have not always aligned with empirical evidence (Maniar, 2007).

Recruitment of Participants

Kushniruk (2005a) used 10 participants to reach saturation. This study extended the former and was expected to require a similar number of participants to ensure saturation was reached. Each participant was a medical doctor or nurse practitioner, i.e., a clinician who read, interpreted and formed clinical impressions on patients' medical histories. This study required participants be familiar with mobile device technology. A brief

telephone question and answer interview was used to assess the prospective candidate's suitability for participation in the study.

To establish a list of participants, the researcher composed a recruitment letter explaining the nature of the study and information on how to contact the researcher. The researcher then, in chronological order: (1) contacted local clinics in person, or by telephone or email using publicly available information, (2) contacted colleagues in-person, (3) advertised for other candidates using the Department listserv using the recruitment letter, and upon receiving permission, (5) posted the recruitment letter on notice boards around the University. To increase the number of participants using a snowball method, each candidate was asked (1) to pass on the recruitment letter to any other potential candidates and (2) for permission to post a copy of the recruitment letter on their office noticeboard if one was available.

Randomization of Participants

To reduce the number of participants needed to gather data, this experiment a within-subjects design was used to evaluate the effect of screen size (Greenwald, 1976). This required that each participant used both devices. To use both devices, the participant required a pair of patient-data cases, one for each device in order to avoid becoming primed with the data from case, before accessing the second case helped to maintain a novel experience for the participant on both devices. Therefore, this study required two cases, designated: *a* and *b*. The devices, *l* and *s* for "large" and "small" screen size devices, i.e., the iPhone and iPad.

To avoid bias potentially caused by the participant always being shown one of the two devices first, the cases and devices were randomized into a block using a permuted-block

randomization process (Matts & Lachin, 1988). Similarly, to avoid any bias caused by the order of the case presented, the cases were also included in the permutation (Rubin & Chisnell, 2008).

For the large screen L there were two permutations: (1) the large screen l used with case a , and, (2) the large screen l used with case b : and together with similar permutations for the small screen, the resulting blocks are shown in Figure 3:

$$\begin{array}{cc} \{l, a\} & \{s, a\} \\ \{l, b\} & \{s, b\} \end{array}$$

Figure 3: Randomizing Screen Sizes with Cases

Permuting l and s together in such a way that neither the case nor the device was ever repeated in a pair, gave the following block of pairs describing what participant P used, i.e., this may be expanded to read, “Participant One used the large-screen device with case a , followed by, the small-screen device with case b ; and, the following participant will use the large-screen device with case b , followed by, the small-screen device with case a . This was shown in the following blocks labelled P_1 for participant one and P_2 for participant two as seen in Figure 4:

$$\begin{array}{l} P_1: \{l, a\} \\ \quad \{s, b\} \\ \\ P_2: \{l, b\} \\ \quad \{s, a\} \end{array}$$

Figure 4: Randomizing Over a Pair of Participants

To avoid any bias caused by the participants always using the large device first the block above was balanced by alternating the devices presented first giving the four-participant final block shown in Figure 5:

$$P_1: \begin{Bmatrix} l, a \\ s, b \end{Bmatrix} \quad P_3: \begin{Bmatrix} s, b \\ l, a \end{Bmatrix}$$

$$P_2: \begin{Bmatrix} l, b \\ s, a \end{Bmatrix} \quad P_4: \begin{Bmatrix} s, a \\ l, b \end{Bmatrix}$$

Figure 5: A Randomized Block of Four Participants

Given that there were four participants per block, the total number of participants in the study was ideally a multiple of four to ensure that there would be equal counts of participants who:

- used the device x (the iPad[®]) first
- used the device y (the iPhone[®]) first
- worked with case a first
- worked with case b first

and, equal numbers of devices that:

- displayed case a first
- displayed case b first

This was a non-stratified permuted-block design; if the total number of participants was a multiple of the block size, four, then an equal number of devices and cases were seen by participants, i.e., a balanced set of variables, was guaranteed and was the reason for a final sample size being a multiple of four (Matts & Lachin, 1988).

Participants enumerated and assigned devices and cases beforehand to minimise confusion during each interview. The roster is shown in Table 1.

Table 1: The final block for groups of four participants

<i>Participant</i>	<i>Device</i>	<i>Case</i>
One	iPhone®	<i>a</i>
	iPad®	<i>b</i>
Two	iPhone®	<i>b</i>
	iPad®	<i>a</i>
Three	iPad®	<i>a</i>
	iPhone®	<i>b</i>
Four	iPad®	<i>b</i>
	iPhone®	<i>a</i>

Materials

The materials needed for this study were an iPhone®, iPad®, the software app, a laptop computer, screen recording software, an external computer storage device and an external video camera. The iPhone® model 4, (115.2 x 58.6 mm) and the iPad® model 3 (241 x 186 mm) were chosen for this study. They were the earliest versions of each device to support airplay mirroring, i.e., the ability to wirelessly send screen video to a suitable host. Both devices were running the most recent, available operating system.

VitalHub®, a Toronto-based software company made an app called *VitalHub® Chart*, referred to through this thesis as “the app”, which ran on a mobile device, connected to an

EMR and provided a mobile device view of patient data to the clinician (VH Chart. 2018). Using its mobile graphical user interface, the clinician was asked to view the patient's vitals, orders, medications, allergies, reports, and medical imaging. The app had a training mode that utilized a small, internal database of fictitious patient data.

VitalHub[®] Chart was the software app used by the participants in the study. This is because it operates on both iPhone[®], and iPad[®] devices. For a more through view of this application, its visual appearance, and operation: please refer to Appendix 7.

A laptop was used to display and record the participant at work, their mobile device screen, and their voice. A live image of the iPhone[®] or iPad[®] screen was mirrored on the laptop's screen using iOS screen-sharing (airplay) technology. The laptop screen was captured with screen recording software along with the laptop's internal microphone as an additional audio source. Please refer to Figure 2 for an illustration of all the parts of the setup used to record the session. This configuration (explained in a step-by-step fashion in Appendix 7) resulted in single video datafile containing multiple items of synchronised media. One such file per participant was captured on the external hard drive stored in a secure location when not in use.

Cases

During the study, participants used the app to find and review two clinical cases. The app contained a sample set of patients and medical data which was accessed without a connection to a live EMR. The data was fictitious, static and could not have been edited on the device. The medical information contained within the training database could not be used to form a complete impression of a patient's wellbeing; for example, many patients had identical referral and admission reports that should have raised suspicion

amongst the participants. For this reason, the two cases from the sample provided by the app were selected on the basis that they were located in two different lists within the app, see Figure 6 for an example of a pair of patient cases and notice how the data differs between the two patients. The participants needed to navigate to, or search for each of them independently.

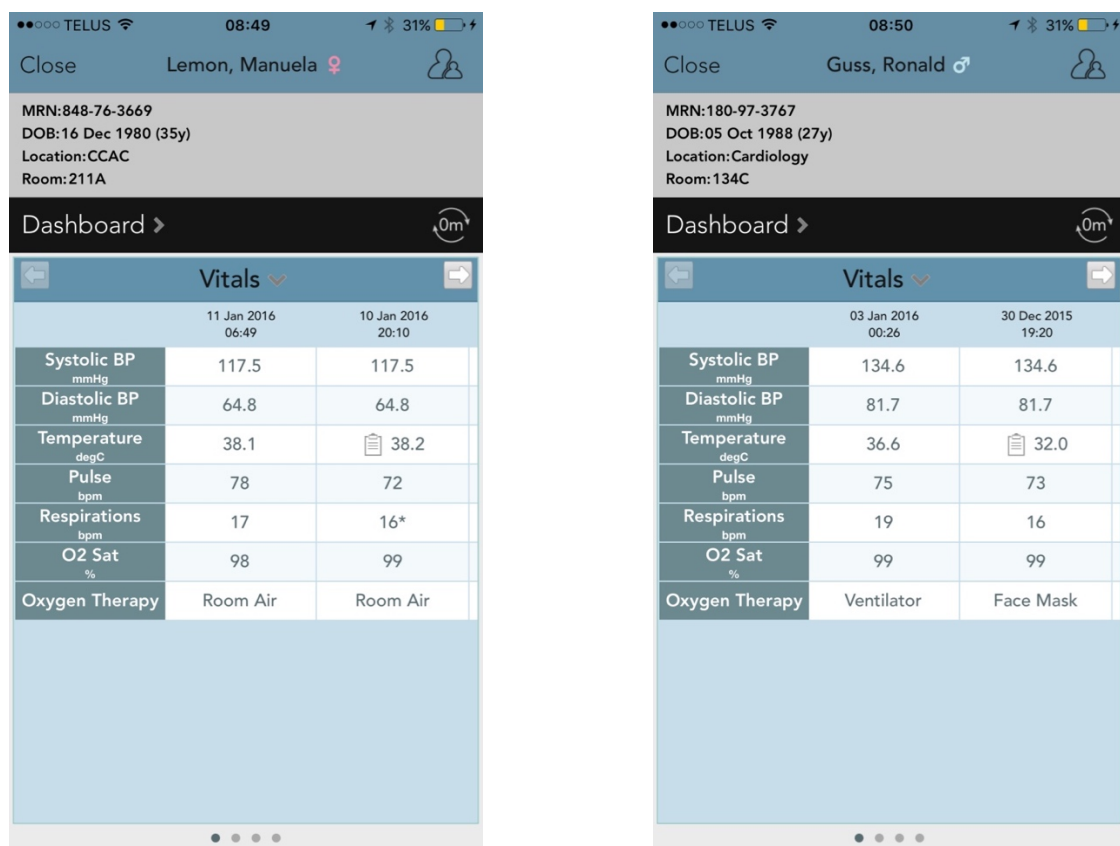


Figure 6: Screenshot of patients A and B

Procedure

Demographics, instruction and training. The user was given a Demographics questionnaire and asked to complete it prior to beginning the study. Gathering data for this study was contingent on the participant speaking aloud while they were doing the tasks. Participants needed prompting when necessary. The initial tutorial part of the session was used to introduce and practice this behaviour.

The app had a training mode that quickly guided the user through its major sections. The participant will be prompted to complete this training and encouraged to think-aloud while doing so. Having presented the device to the participant the following narrative text was read to them:

“The first thing we’ll do is introduce the app and walk through the tutorial pages. Use the iPad®, open the app and tap on the help button along the bottom. Please read each paragraph aloud and swipe the images to navigate through the tutorial. When you’re finished, we’ll move on to using the app. Please remember to read the tutorial aloud and to think-aloud any comments or thoughts that may occur.”

Before proceeding, the participant was allowed time to finish the tutorial. Before starting the session, the participants were reminded of the need for thinking aloud. When ready, the participants were prompted to locate the first case:

“Please use the app to locate Patient [patient name] and begin reviewing the patient’s data. Please continue to think-aloud during your investigation and be prepared to summarise your impression of this patient’s health when you are complete. What do you think is wrong with them?”

The purpose of asking the participant to form an impression and to summarise the patient's health was intended to provide a goal so as to have encouraged the participant to engage in the assessment process permutation (Rubin & Chisnell, 2008).

“The first task is to examine the watch-list on the home page. Using the ‘Vitals’ information please vocalise your impressions of what you read about the patient in case *a*. If you accidentally tap away from this area, please return to it and continue until you have reviewed all the patients in the list.”

When necessary, the participant was reminded to think aloud in order to capture their thoughts in the video recording for later analysis.

“From the home screen, please locate the patient and review their full charts. Please verbalise your impression and any finding and don't forget to think aloud while you are working.”

Post-case interview. Following the first device session, the participant was given a sheet of paper with the following five statements and each with a nine-point Likert scale that was weighted such that at the low end of the scale, 1, was labelled Strongly Disagree and the high end, 9 was labelled Strongly agree:

1. It is easy to use the app.
2. It is easy to navigate the app.
3. It is easy to read the information presented.
4. It is easy to use the device.
5. It is easy to understand the information presented.

Immediately following the Likert questionnaire, the participant was interviewed in a conversational manner using the following semi-structured questions, read out loud:

1. Did you find this app easy to use for this case? Why or why not?

2. Did you find the app easy to navigate for this case? Why or why not?
3. Did you find the information easy to read for this case? Why or why not?
4. Did you find have any problems using the app? Please discuss?

The participant was prompted for further comment and asked any follow-up questions that were necessary. The conversation continued until the subject naturally came to an end (Rubin & Chisnell, 2008).

The participant was moved on to the second device and the actions described above were repeated.

Final interview. After the participants used both devices, completed both questionnaires and completed both interviews, then the final questionnaire was started. The participant was given another sheet of paper with the following statements each with a nine-point Likert scale on which to provide an answer. The Likert scale used had nine points and was weighted such that at the low end of the scale, 1, was labelled Strongly Disagree and the high end, 9 was labelled Strongly agree:

1. The iPad[®] is easy to use in clinical practice.
2. I would use the iPhone[®] with this app in clinical practice.
3. I would use the iPad[®] with this app in clinical practice.
4. The iPad[®] with this app is safe in clinical practice.
5. The iPhone[®] with this app is safe in clinical practice.
6. There are advantages with iPhone[®] when using this app.
7. There are advantages with iPad[®] when using this app.
8. The iPhone[®] easy to use in clinical practice.
9. It is easy to use this app with the iPhone[®].

10. I am confident I can retrieve the information I need using the iPhone®
11. The iPhone® is too small for clinical use.
12. It is easy to use this app with the iPad®.
13. Using the iPad® when working with patients is unsafe.
14. Using the iPhone® when working with patients is unsafe.
15. I would avoid using the iPad® in clinical practice.
16. I would avoid using the iPhone® in clinical practice.
17. The iPad® is too large for clinical use.
18. I am confident I can retrieve the information I need using the iPad®.
19. This app is not safe for use in a clinical setting.
20. This app is safe for use in clinical team meetings.

The final question on the form was: “Which device did you like better?” with only “iPad®” and “iPhone®” to choose from.

When the participant was done, the final interview began by reading the first of the following four questions which were also delivered in a conversational manner with prompting and follow-up probes when needed:

1. Would you say that the iPad® or iPhone® are difficult to use in clinical practice?
Why or why not?
2. What are your thoughts about patient safety while using these devices in clinical practice? Please discuss.
3. Are you confident that you would be able to retrieve and interpret patient data by using these devices? Please discuss.

4. Are the screen sizes appropriate to use in clinical practice? Why or why not?

A copy of the printed version of the questionnaires used in this study can be found in Appendix 10.

Data Collection

Data for this study was collected in the form of Likert-based questionnaires, audio recording of the interviews and audio-visual recordings of each participant using the devices. Each participant was given a number, the paper documents were marked with this number and the media files were named with it to maintain organisation.

Analysis of the Questionnaires

Demographics questionnaire responses were transferred to a spreadsheet and processed to compute mean, mode and median values for each question. Likert responses were tabulated from all questionnaires and organised on a spreadsheet. Likert scores were summed across all participants and the mean, mode and median scores were computed for each question. Results were compared across the iPad® and iPhone® to reveal differences between the responses to the same question on both devices. The results were plotted in several visualizations to gain insight.

Analysis of the Transcripts

The audio from the recording sessions were transcribed with the timecode to indicate at what time certain things were said. This data was saved into plain text files formatted in manner similar to that of a dramatic screenplay because that particular formatting style

lends itself to the need of separating spoken content is clearly separated from comments about the content (Carless, 2004).

Content analysis was used on the transcribed interview sessions. The scheme used was based on the Think Aloud protocol analysis guide found in a paper by (Kushniruk & Borycki, 2015). This is a method of annotating a transcript with metadata. This work was done in two passes, once for the transcript and once again while viewing the screen recordings.

This analysis methodology uses a system of attributes to indicate facts about the content. These facts are written in the form of attributes, typically one or more phrases written near the relevant transcribed text. These attributes, called codes, are classified into three top level groups: “Usability”, “Usefulness, and “Safety/Error” in which all other codes are a subset. Usability codes were used when the participants did or said something that related to a problem with app’s user interface. Similarly, Usefulness codes, signified the app’s value as a tool, and Safety/Error codes were used when the participant made a mistake with the device. The codes used in this study were based on the qualitative codes provided in Kushniruk and Borycki (2015) and expanded when necessary.

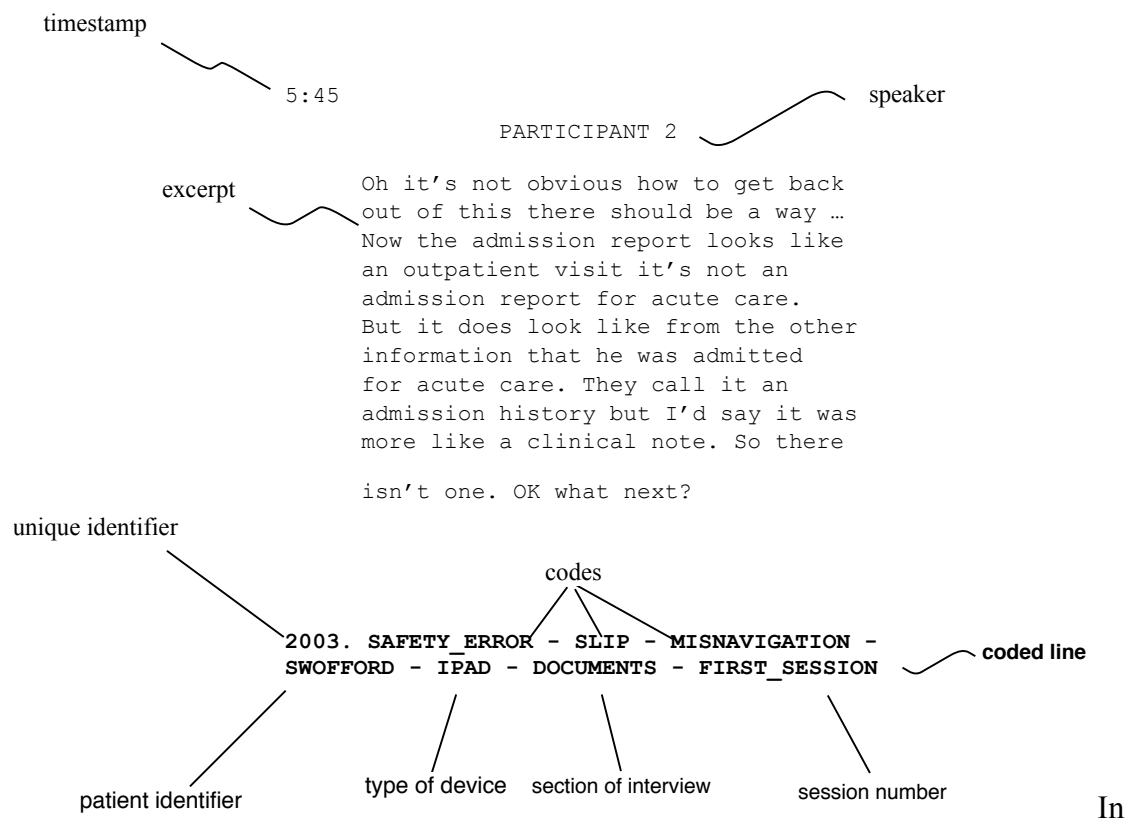


Figure 7: Anatomy of a coded excerpt

In Figure 7, at the top-left, there is a time stamp indicating that this text appeared at five minutes and forty-five seconds into the recording. Near this time this participant made a navigation error and inadvertently exited the patient record while attempting to navigate to the previous screen. The evidence was visible on the screen recording and implied on the transcript. This was a minor error that was caught by the participant before continuing with their task. This incident was coded with a code-category of “SLIP” and, a code-type of “MISNAVIGATION” that can be seen in the coding line below the transcribed text in Figure 7. Other utility codes were also present in this line. Read from left to right, they were: SWOFFORD, the fictitious patient’s name, IPAD®, the device the patient was using, DOCUMENTS, the section of the app where the incident occurred, and

lastly, `FIRST_SESSION`, indicating which of two sessions the participant was completing. Additionally, the leftmost code, 2005, is a unique identifier for this code, i.e., participant 2, code 5, it was used to refer back to this line when needed during the analysis.

The code-lines had a strict format devised by the researcher to facilitate being processed by a computer. To that end: all codes in the transcripts started with a unique identifier and were written using the following template (items in brackets were optional):

- ID – CATEGORY – CODE-NAME – [CODE NAME ...] – PATIENT – DEVICE – SECTION-OR-CONTEXT – [COMMENT] – POSITIVE-OR-NEGATIVE – [***] – [IMPORTANT]

The following guide, devised by the researcher, was followed when coding the transcripts:

- Each item in the code-line was filled in once with the exception of CODE-NAME which could appear many times; in other words, each line had only one category but could have many code-names
- At a minimum, use all non-optional parts of the template
- Use one line per coding incident
- Comments are always tagged as being either positive or negative
- Comments are always tagged with the code name
- All codes are considered to be problems unless tagged positive
- Three literal asterisks denoted an especially pertinent comment
- Novel or interesting ideas were flagged with IMPORTANT.

For a complete account of the Usability, Usefulness, and Safety codes used and their meanings, see below for Table 2, Table 3, and Table 4 respectively. For a complete

example of a coded transcript see Appendix 13:.

Table 2: Usability Coding Dictionary

This table lists the *Usability* codes used to catalogue incidents in the transcripts and screen-recorded data. These codes follow the scheme published in (Kushniruk & Borycki, 2015). Codes listed in italics are extensions devised by the author.

<i>Code</i>	<i>Incident was a:</i>
NAVIGATION	Problem moving through a system or user interface
CONSISTENCY	Problem due to a lack of consistency in the user interface
TERMINOLOGY	Misunderstanding about language or labels
COLOUR	Problem or comment about the colours in the UI
EASE_OF_USE	Comment on overall usability of the app
FONT	Comment or problem due to readability
GRAPHICS	Comment or problem with visual item
INSTRUCTIONS	Problem understanding instructions
LAYOUT	Problem due to screen layout
STATUS	Problem understanding the status of the system
WORKFLOW	Negative impact of workflow on app use
<i>SCREENSIZE</i>	<i>Problem or comment about the size of the display</i>
<i>ORGANISATION</i>	<i>Problem or comment on the structure of the app</i>
<i>LEARNING_CURVE</i>	<i>Problem or comment on the app's complexity</i>
<i>COMPREHENSION</i>	<i>Misunderstanding what they saw in the UI</i>
<i>CONTROLS</i>	<i>Misunderstanding how to use the UI</i>
<i>FAMILIARITY</i>	<i>Comment about familiarity of the app</i>

Table 3: Usefulness Coding Dictionary

This table lists the *Usefulness* codes used to catalogue incidents in the transcripts and screen-recorded data. These codes follow the scheme published in (Kushniruk & Borycki, 2015). Codes listed in italics are extensions devised by the author.

<i>Code</i>	<i>Participant's comment expressed:</i>
APPLICABILITY	The information presented was not applicable to real healthcare practice or cases encountered
CORRECTNESS	The information presented was incorrect or inaccurate
RELEVANCE	The information presented is not relevant to the task at hand
TIMELINESS	Problem or comment about the speed of the user interface
IMPACT	Unexpected impact of the system on work activities
<i>PORTABILITY</i>	<i>Comment about the ease of carrying the device</i>

Table 4: Safety and Error Coding Dictionary

This table lists the *Safety* related codes used to catalogue incidents in the transcripts and screen-recorded data. These codes follow the scheme published in (Kushniruk & Borycki, 2015).

<i>Code</i>	<i>Participant's behaviour revealed that:</i>
SLIP	A mistake was made, noticed and corrected before continuing
MISTAKE	A mistake was made, noticed and not corrected before continuing.
WORKAROUND	Coded when the user is not using the approach to carrying out work that is recommended by the app.

Analysis of Data Processing

The lines of coded information were collected for each transcript and compiled into a spreadsheet. Each participant's demographics data was also included adjacent to the corresponding transcript data. For each line in the spreadsheet, the codes and demographics were split in such a way that each part, separated by a dash, resided in its own column. The columns contained one field each, for example, one column contained the participants' age range, another contained the code-name, and yet another the code category. In this way each column contained all the participants' data for a field. The source row was searched and automatically split using built-in spreadsheet functions. The resulting split-up data in the worksheet was converted into a table for ease of manipulation. The resulting table was, in turn, converted into a pivot table to enable the analysis to be done by filtering and cross-referencing fields with one another.

Summary

This section outlined the recruitment of participants, and their randomization. This task was much more difficult than expected. Physicians in this geographical area are apparently in short supply and are difficult to meet due to their stated busy schedules.

This anecdotal evidence is mentioned as it will be used later to discuss alternative recruitment strategies. Participants were given paper-based forms to complete and took part in a randomised simulation of a chart review whilst thinking aloud and being recorded along with a screen recording of the display they were using. Data was processed according to methods used in previous study by Kushniruk and Borycki (2015) and the Usability Guide by Rubin and Chisnell (2008).

Ethics Approval

The University of Victoria requires ethics approval for research involving human subjects. An application for ethics approval was submitted to and accepted by the University of Victoria's Human Research Ethics Board office prior to commencing data collection. Refer to Appendix 1: for the approval documentation.

Chapter 5: Results

Introduction

This section of the thesis provides an overview of the analysis. Beginning with a summary of the data processing, followed by a review of the participant demographics, this chapter provides the results in three parts: iPhone®, iPad®, and a comparison between the two.

The Software Usability Measurement Inventory combined with the Low-cost Rapid Usability Engineering were used to evaluate one group of participants. Screen capture-based usability data, Cognitive Task Analysis and usability heuristics were used to aid in analyzing the video data (Kushniruk, Patel, & Cimino, 1997).

Participant Demographics

Nine participants took part in the study. From data gathered in the demographics questionnaire, it was known that all, but two participants were over 44 years of age. The median age range was 45 to 64, and the largest group was 65 to 84 years of age. This information is shown in Table 8 and in greater detail in Figure 9.

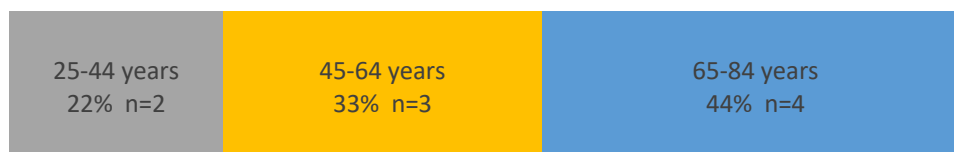


Figure 8: Participants' Age Ranges

Some participants, 44%, were new to medical apps and none were new to mobile devices. Most, 89%, had had at least five years of experience with mobile devices. All participants that responded indicated that they had ten years or more of IT experience.

The eldest participant had used both mobile devices with great ease. This participant, a physician, had the most experience with computing, mobile devices, and medical apps – and was currently in active family practice.

In general, a typical participant in this study had: less experience with medical apps, some experience with mobile devices, and more experience with medical practice and computing.

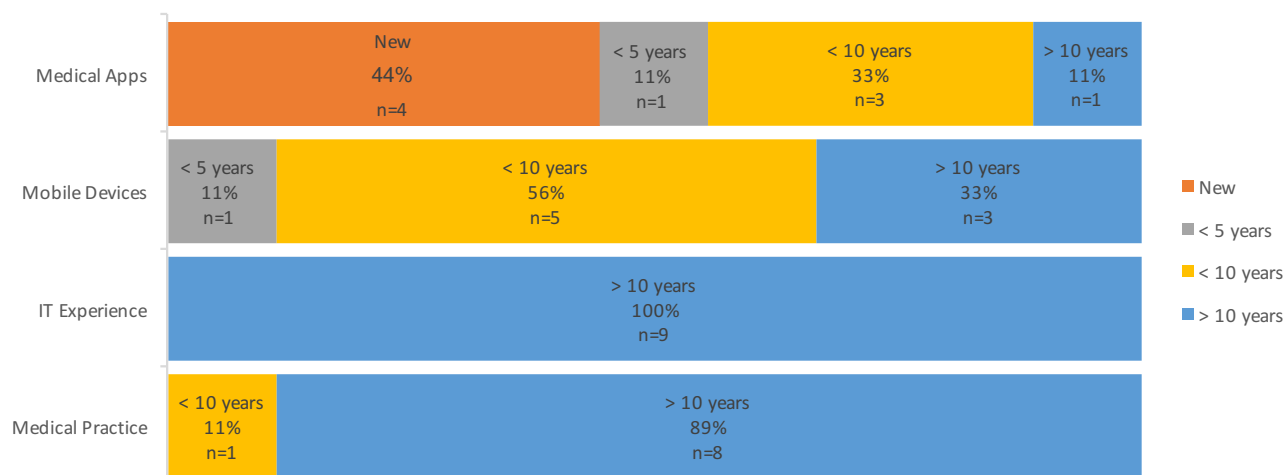


Figure 9: Participants' Experience

Notes

In the next sections, the results from the coded transcriptions are presented. For readability and clarity, the participants' words are presented as follows:

Participant 1:

Temp is low taken under her armpit,
pulse is good. If she has an armpit,
she must have an arm.

Participant 2:

Oops! I'm stuck

The horizontal line which separates two or more quotations, signifies that each quotation is from a different participant. Two types of finding emerged during coding: 1: actual problems participants encountered during the session, and, 2: comments made by the participants.

The iPhone®

This section of the chapter describes the results from participants' interaction with the app using the iPhone®. The task given to the participant was: "Please review the data and summarize the patient's health and status." Participants were assigned a specific patient to review.

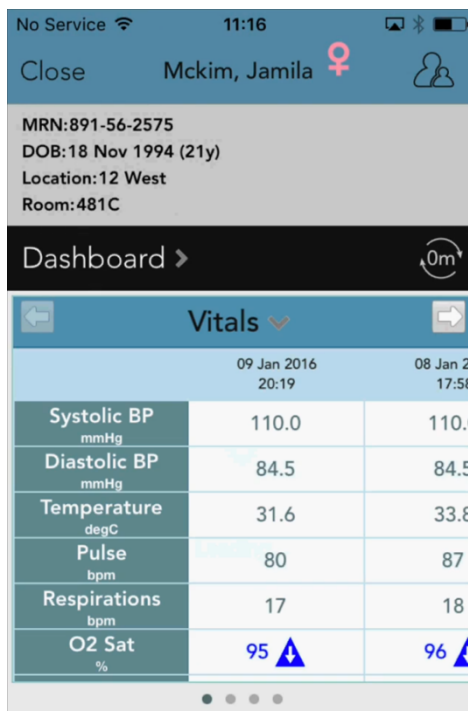


Figure 10: The iPhone® app main dashboard participants' view

This figure shows what the participant saw after selecting a patient. The main navigation control was the word 'Dashboard' in mid-screen. Tapping the word 'Dashboard' opened a list of navigable items.

A total of 105 problems were recorded and coded in the participants' actions and video recordings. Participants made 87 positive, 79 negative and, 9 neutral comments on the iPhone®. These were broken down into categories shown in Table 5 below.

Table 5: Problem and comment scores by code for the iPhone®

Please refer to Table 2, 4 and 5 on page 80 for definitions of the Codes.

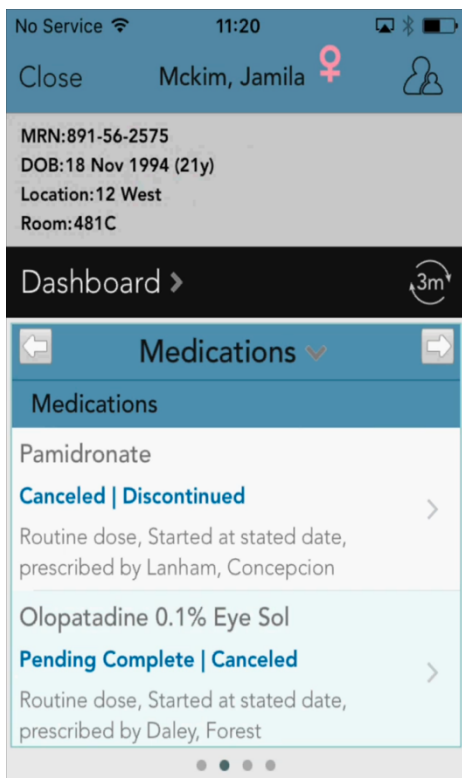
Codes	Problems	Positive Comments	Neutral Comments	Negative Comments
Error				
Mistake	2	1	2	0
Slip	11	0	0	0
Workaround	3	0	2	0
	16	1	4	0
Usability				
Colour	1	1	0	0
Comprehension	15	0	0	1
Consistency	15	1	0	2
Controls	5	7	0	10
Ease of Use	3	9	1	9
Familiarity	1	1	0	1
Font	2	8	1	3
Graphics	4	8	0	0
Instructions	1	0	0	0
Layout	5	3	0	5
Learning Curve	0	2	1	0
Navigation	13	9	1	12
Organization	5	5	0	3
Screen Size	3	4	0	17
Status	1	0	0	0
Terminology	1	0	0	0
Workflow	3	0	0	5
	58	68	4	78
Usefulness				
Applicability	8	5	1	3
Correctness	14	2	0	3
Familiarity	1	1	0	0
Impact	4	1	0	7
Portability	0	8	0	1
Relevance	3	1	0	0
Timeliness	1	0	0	0
	31	18	1	1
Total	105	87	9	79

iPhone® usability problems. In this section the problems participants had in using the iPhone® to complete the assigned tasks are described. Usability problems had three main causes: Comprehension (19%), Consistency (19%) and Navigation (16%). Please refer to Table 2: Usability Coding Dictionary on page 80 for definitions of the foregoing terms.

Comprehension of terminology. The app is intended for use in a hospital setting. Not all the terminology used in the app was familiar to all participants because not all participants were hospital physicians:

Participant 6:
Some of the results are a little
unfamiliar to me, I'm not an acute
physician.

Other participants were confused by the wording used in the app; for example, when a user interface element appeared to be, but was not actually, a control. At least two participants were confused by the wording and appearance of text, illustrated in Figure 10, in the Orders view of the app:



Participant 2:
 "Pending Complete Canceled" what does that mean? Are there two sections? I don't know what it means at all.
 Okay. These are three medications that the patient is getting ... the first one has been started. [This patient] is getting medication ... last updated ... Verified? Is this a verification or does this apply?

Participant 1:
 When I see the word it is in the past but [it looks like a control] and I want to [tap] it to cancel the order or discontinue the order.

Figure 11: iPhone® Orders View

Some words shown in the UI looked like controls to some participants who noticed the appearance of the word 'Cancelled.' Some participants interpreted it as a being tappable, i.e., that one would tap it to change its state or learn more about the fact.

The medication screen showed a heading which appeared to be a control. The participant was unsure as to whether the word 'Verified' at the bottom of the page was an indication that this item was actually verified, or was a control they could use to set, or perhaps verify, this information.

Graphics and Layout problems contributed to the lack of comprehension. The time-series values used for many vitals on the app were often identical. Tapping on a name of the value, for example, BSA, caused a chart to appear which displayed as a flat line due to the series of identical values. Examples of these charts were shown in Figure 12. As a

result, the chart confused the few participants who discovered the graphical feature as seen following participant's statement and in Figure 12:

Participant 1:
 OK. Oops, I've pressed something and I've ended up at [a graph of values] of 'BSA' [and] I don't know what this is.

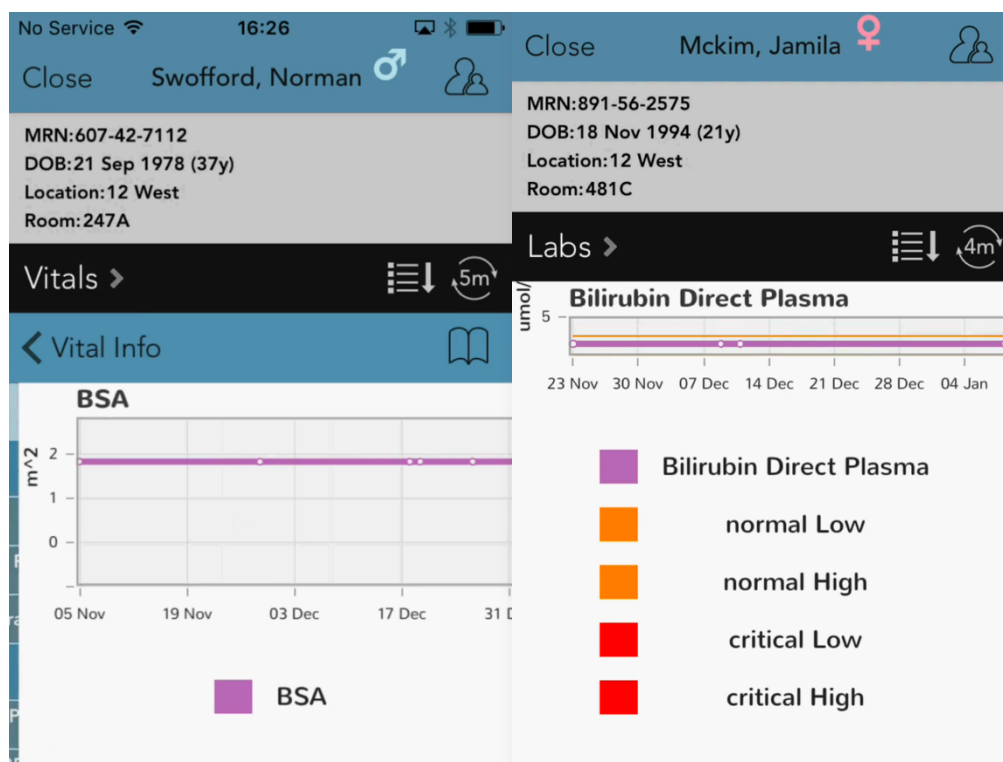


Figure 12: Charts on the iPhone®

Charts on the iPhone® were rendered as flat lines from non-variant time-series data, caused confusion in some participants. Even if the right-hand chart showed non-variant data, it would be difficult to read because there was so little space on the Y-axis.

Some values on lab results had no units. The following participant saw a numerical value but no units to qualify the value:

Participant 7:
Protein Total Plasma

[swiping, tapping]

I'm interested to know if anything's abnormal. Glucose urine drip ... oh critical ... doesn't say [what the units are, but] it seems to be saying that it's a mistake

Participant 4:
Let's see what's in urinalysis.

[reads some values out loud].

'Dip' as opposed to 'drip'. Volume Daily Urine, 2.6 ... 2.6 what? Litres? That's a lot.

Other instances of comprehension problems were found due to the artificial quality of the demonstration data in the app. Incorrect information on some records made the task of assessing one female patient's record difficult:

Participant 4:
... and there are prostate notes in here and it's marked "men only", well surprise, surprise

More than one participant found the app's medical information to be incomprehensible. The following was a detailed example of one participant's experience as they attempted to interpret the data:

Participant 4:
OK so we have normal, diastolic, a remarkably low temperature. So she's come in hypothermic? Or she's post-cardiac or brain where you reduce the temperature so that the metabolism

doesn't go ... so you don't get expansion and that kind of nasty stuff. ... ok so O2 saturation ninety-five that's not too bad ... if she's that cold you'd have thought the pulse would have dropped down quite a bit and you'd expect the respirations to be slower than that too... I would say [reads more stats] well they seem to have tried everything with this sucker. That's in November 23rd ... ah wait a minute ... the temperature is 38 in December and drops to 33 in January that's very odd. Thirty-nine point six is high, but the pulse isn't particularly high though the BP seems to be absolutely static, the pulse was low in November but other than that it's 82, 83,

[continues reading stats]

... and yet she was on a ventilator when it was a little bit high but her BP was spot on normal.

I don't understand this patient.

Another participant expressed surprise at not being told where blood pressure was measured on the patient. After a briefly reading the data, the participant seemed to mistrust the quantity and scope of medical problems diagnosed on a single patient:

Participant 1:

Blood pressure's fine, but they didn't say why they took blood pressure on her right leg, right arm? Left leg? What on earth were they doing? She has some weird vascular disease, I need to find out...section two...I guess...so I'll look at the beginning, scans. One cancer follow-up. Need to get off vitals and go somewhere else to see what the hell is going on. Let's look at Documents and see. [reads content]

Lung cancer? She seems awfully young to have Lung cancer.

Participants were asked to rate their response to the statement: “I am confident I can retrieve the information I need using the app.” The median Likert score (6) indicated the participants as a group agreed they could retrieve the clinical data they needed but they did not agree strongly. Some comments noted that the app did not allow them to view more than one set of data on the iPhone®’s small user interface. Ironically, because apps like this purport a reduction of paper use, one participant considered that paper would aid them when using the app:

Participant 5:

Oh probably. I think so. I think that I would find doing this on the iPhone® is that I'd have a piece of paper next to it and I'd end up jotting notes as I went.

Consistency of the user interface. Some users were irritated by the user interface inconsistencies on the iPhone®. One participant who had reviewed the Orders section, moved on to the Labs section and commented:

Participant 1:

I'm going to check the content of the first order. Why is the design of this different? I prefer a [consistent] design. Because I set my mind and now I have to change it.

The iPhone® version of the app had no dedicated section for viewing diagnostic imaging. Some imaging was available in the Reports section, but was not labelled as such. As a result, some participants were unable to find the medical imaging; other participants persevered, with minor prompting:

Participant 2:

Is [the imaging linked] from the dashboard? Let's see ... well it's not labs we've been in there, it's not Documents, we've been in there ... can't be vitals, unless it's under ...

The researcher, who was aware that the imaging reports were in "Documents", prompted: "Would it be under Documents?" and the participant gave the following response:

Participant 4:

Well we were just in there and didn't see any. Let's have another look. Under "Infection-Tumour" well there's a gallium scan ... a Right "i-ograph" Yeah here we are. So these are under Documents. I found it only by a process of elimination of everything else.

Hierarchical inconsistencies were noticed by one participant. Once noticed by the participant the inconsistencies became a distraction for the participant on subsequent pages as described below:

Participant 9:

You know, it could be OK for the user but with one [condition]: if Vitals, Labs, Orders Documents are in the same hierarchy and they can open the other [sub-] pages then it's ok.

But if the pages have similar names then it makes it a little bit confusing. When I opened Vitals it brought me to a section where there was another page names Vitals with more detail. So it's a little ...
[confusing]

Navigation within the app. Navigation was another other major usability problem on the iPhone® app. The most common navigation error had to do with returning to a

previous view, after having gone forward for more detailed information. There was no 'back' button on the user interface and this was noted by several participants as seen in the following excerpts:

Participant 1:

Oh here's a little button I can push here. Uh-oh, don't know what I did there. That now gave me everything when I pushed that arrow but I was trying to figure out how I'm going to go back to the dashboard. OK back to my orders now that's what I want to see.

Participant 2:

Oops ... Oh I'm stuck in this need [and I] need to get back

[tries to navigate, ends up back to the list of patients]

Participant 3:

I ... uh ... I'm stuck

iPhone® usability was hampered by inconsistent user interface controls and labels such as the back button, however, all participants were able to cope with the problem by devising a cumbersome, but effective, workarounds. The comprehensiveness of the outdated and at times bizarre medical conditions found on the patients in the app seemed to cause participants to lose trust in the information and left them with an unfavourable impression.

iPhone® usefulness problems

The most common Usefulness problems experienced by participants were: Correctness (44%) and Applicability (25%). Please refer to Table 3: Usefulness Coding Dictionary on page 80 for definitions of the foregoing terms.

Correctness of the app data. The app's reports, and other items with addresses, were apparently from Canadian locales. Several participants reported that some of the units and names were not commonly used in Canada and seemed more likely to be from the U.S. As a consequence, the participants did not fully comprehend the patient information presented. The fictitious data in the app appeared to be intended for a U.S. audience. For example, the numbering system was unfamiliar to participants as seen below, could lead to an error:

Participant 2:

I'm going in the labs section. So glucose there's 7.3 and that's ok ... Well it doesn't really give the units but 7.3 is a typical value that we see. They use a different numbering system from what we use here [in Canada].

On reviewing a patient's vitals, some reference values seemed incorrect. For example, one participant reviewed the reference ranges for Body Mass Index, a derived value used to measure the patient as underweight, normal weight, overweight or obese, based on that value {Mei:2002ij}. This participant, a GP, disagreed with the reference ranges shown below:

Participant 7:

Oh it's got height here ... weight, body mass index well there it is, perhaps we missed on the other one [iPad®] well that's jolly good -- quite thin. [the reference ranges are] normal 20.7 critical 27.5 ... well I wouldn't consider it normal at 27.5 I don't agree with that.

Two participants experienced cognitive dissonance, when they reconciled the data in the table views with the data in the Reports section. Due to the artificial quality of patient data, the participants found some errors startling as noted in the following excerpts:

Participant 5:

I pressed close and then went straight back to [the patient] again. Ok let's look at progress reports.

[reads report titles]

Hypertension, diabetes ... What's this? According to the other sections this patient doesn't have any of these things.

Participant 8:

this report is not for this patient it's [for] a male outpatient!

Some other patient data seemed physiologically unlikely. These values may or may not have had plausible causes, but they stood out as extraordinary to these participants as seen in the excerpts:

Participant 8:

Ok this a young [patient] twenty-one years old. Ninth of January ... her temperature is 31.6 degrees which is like ... kinda ... dead. So that's bad.

Participant 6:

Inspired oxygen ... there's 18.6 percent there something wrong with that.

Four participants commented that some information was unorthodox or inappropriate. Although there may have been plausible explanations for the results of these following patients, the participants commented:

Participant 1:

Well I'm looking at the oxygen...
 Blood pressure on the right leg? Well
 that's a bit queer. It's usually the
 left arm or the right arm. ...
 The pressure's okay. Cuff placement on
 her left leg? What the hell?
 Any [woman] of 21 if I asked her to
 pull her trousers down and take the
 blood pressure on her leg, she might
 wonder what I was after.

Participant 2:

I don't know why she's had vitals done
 these last two days usually we'd only
 do that with someone who's acutely ill.
 But someone wanted lab values at 5:20
 in the afternoon so I don't know what's
 going on there and there's things that
 I cannot tell from this necessarily.
 ... It's like [this data is from]
 clinics within the hospital. Like that
 last patient seemed to be inpatient but
 all the reports seemed to be outpatient
 based. Mostly routine follow-up.

One participant commented that the reports did not state the patient's name as in this
 excerpt from the transcript:

Participant 7:

...So I don't see the name of the
 patient on this piece-of-paper report.
 There's no point having a report
 without the patient's name on it. It's
 a mistake

Unidentified reports, strange information, and a sense of artifice are all strong
 impressions the participants were left with after using the training mode of the app for
 half an hour. The data presented to participants was at best not believable, and at worst,
 distracting.

Applicability of the app data. The applicability of the data was another major problem for “Usefulness”. The medications listed in the demo data suggested that they were out-of-date with respect to current medical practice. To one participant, the inclusion of a medication that is not currently used seemed inappropriate:

Participant 3:

[these drugs are] not used any more, it's for migraine. Ergotamine and Caffeine mixed 'Caffergot' was the name of the drug. Almost .. hardly... anyone uses it anymore. There are better medications for migraine.

Details of some reports seemed to contain information irrelevant to the current patient, and in some instances, outpatient data was shown for an inpatient as noted by participant 2:

Participant 2:

So I guess it's a Progress Report even though it says "patient's brother" ... I don't know what that means ... It's like [this data is from] clinics within the hospital. Like that last patient seemed to be inpatient but all the reports seemed to be outpatient based. Mostly routine follow-up.

A recurring comment from participants who were asked to browse through a patient's data and determine the patient's state of health, was a sense of frustration about the lack of a 'handoff' note, that summarized some basic information about the patient. The participants experienced frustration when they had to process all the raw data to determine the patient's health status, and expressed a desire to see a summary page as described by participant 2:

Participant 1:

So the first thing I see is Vitals. What I want to see is "Patient was

admitted to hospital" you know? That's what I want to know...
 I don't want to start with doing their tests ... reading their test and then figuring out what's wrong with them...
 I'm looking for something that will give me some of her history.

This problem may have been exacerbated by the fact that the patient data shown in the demo screens was not a true record of a real patient. Worse, in the second sessions, participants noticed a pattern emerge, i.e., that identical reports were shown for both current and previous patients. This is described by participant below:

Participant 1:

Wow [this record has] a lot of stuff here. Let's look at the progress reports.
 ... cardiac ... diabetes ... hey wait a minute, this is just like the guy we did earlier ... that doesn't seem right.

The participants experienced usefulness problems almost entirely due to the correctness and applicability of the data presented. The most problematic of these problems is caused by mixed locales, i.e., patients appear to have Canadian addresses but units for lab results are not consistent with those used in Canadian healthcare. This is a potential path to error and is a safety concern.

iPhone® error and safety problems

Most problems in the Safety-Error category which occurred on the iPhone® (i.e., 64%) were coded as "Slips". There were two actual errors. Please refer to Table 4: Safety and Error Coding Dictionary on page 81 for definitions of these codes. One problem occurred in the Allergies section. There appeared to be data missing from the view. With

a little persistence, the participant was able to enter an editing view, whereby the full data was available (see Figure 13 below).

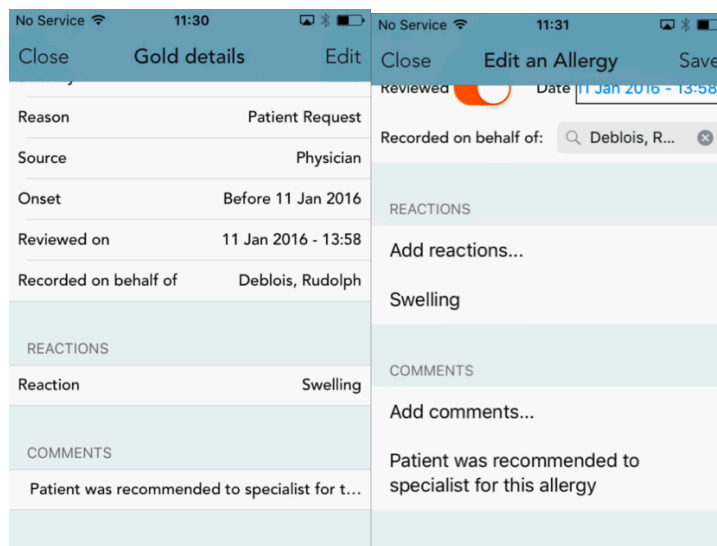


Figure 13: Edit mode on a record revealed

Edit mode on a record revealed more information not visible on the read-only view. Notice that the text under 'comments' was fully displayed only in the edit mode

A “Slip” occurred when a participant made, and then immediately corrected, an error while using the app. Of the “Slips”, seven (i.e., 63% of iPhone® slips) were caused when the participant wanted to “go back” to a previous view, and instead, tapped a ‘close’ button. This caused an exit to the list of patients. On the iPhone®, the user interface did not provide an intuitive way to change the view; for example, from a patient’s Labs, to a patient’s Medications. Navigation was easier from the patient Dashboard, but the Dashboard was not visible from the sub-pages and again, there existed no ‘back’ button as described in the following excerpt:

Participant 2:
 Oops. It’s hard to back up. I meant to mention that on the [iPad®]. So I don’t

really know how to get to the next section [without exiting first].

This problem occurred so frequently, that many of the participants used it as a workaround. Participants would, on occasions when they were deep in some nested view, deliberately quit the patient view. As a result, they could immediately open the same patient record to restart navigation from the initial Dashboard page. When coding the video transcript, if the workaround was done deliberately, the act was coded as a Workaround and not a Slip.

The two actual Errors, i.e., mistakes that were not caught and recovered, participants misinterpreted an icon, and a comment from one participant who, when viewing a graph, had difficulty interpreting the lines due to vision issues, see below:

Participant 4:

Oh, they're giving you colours so you can see high and low values. But I'm colour-blind so this is absolutely bloody useless to me. Presumably the bottom two are red

[they were red]

...orange or a lighter red

[they were orange]

...and the top one is green?

[it was purple]

These issues, though minor, highlight an area that ought to have been accommodated for in a healthcare app. The inconsistent use of the familiar “back button” control, i.e., the presence of a back button in some parts of the UI, the lack of one in other parts in the UI, and the presence of a button in the familiar location that was not actually a back button, was the most consistently frustrating cause of participant slips.

iPhone® comments. Most comments were made by the participants during the interview part of the session. Comments were also made during the part of the session and the codes indicate within which part the comment was stated. In this section, the results of comments made throughout the session will be presented.

On the iPhone® there were a total of 86 negative comments, and 84 positive comments. These comments were categorized and coded using the same scheme and categories of usability, usefulness and error.

iPhone® usability comments. Of the negative comments, the top four categories were: screen size (25%), navigation (18%), controls (15%), and ease-of-use (13%). The positive comments were in these categories: Ease-of-use (16%), Navigation (16%), Font (14%), Graphics (14%), Controls (12%), and others (28%).

Screen size. Participants were negative in their comments about screen-size. Below are excerpts from the transcripts that suggest this:

Participant 1:

[The screen is] **too small** and the speed at which can use a mouse and a decent size screen would be infinitely better...

I think that I would find doing this on the iPhone® is that I'd have a piece of paper next to it and **I'd end up jotting notes** as I went.

Participant 9:

The iPhone® was **absolutely too small**, not enough space for all that information and the fonts were too small ... depending on the demographic of people using them the average age of people who use them in hospital not all

of them are able ... (laughter) ...
 some of them will need glasses

Participant 2:

Not as easy. It **felt too small**, yeah. I think the app itself is ok but the [iPhone®] itself was too small.

Participant 6:

[The iPhone® is] a bit finicky, the screen size is **too small**, if you inadvertently click a thing it's so small. and also, when it comes to looking at the patient's vitals, you would like to know the trend so it's better to have all the results in one screen and it's better to have a snapshot and not swiping.

Navigation of the app within the device. Negative comments on Workflow concerning the dexterity needed to interact with the user interface. The need to physically navigate in order to discover the data on the device, was perceived to have a time penalty as shown below:

Participant 1

Personally, I want a mouse that goes directly where I want. I'm looking at the time I spend getting information from people. I've got ten minutes [with a patient], basically.

...
 Opening and closing, opening and closing everything and that takes time. I set my goals. I start here at 9:30AM and if I have to wait two minutes to see the next patient in the next room, and you multiply that by eight hours, it takes time.

Controls and ease-of-use of the app-device pair. Viewing a PDF report on the iPhone® screen that was impossible to read without zooming and panning to scroll parts of the document was a participant concern. When a participant was asked to describe the need

to zoom and pan, the reply was complacent in tone and the participant recognized that the need to navigate was problematic:

Participant 2:

That makes it difficult, but that's just being aware that that can happen when you're using your finger. So...that can happen when you use the mouse...but it does potentially lead to missing information.

Other participants, however, liked the user interface because it permitted the user to expand and shrink a document. This was likely due to the effect of familiarity with this user interface feature as seen in the following participant comments:

Participant 4:

Given the [small] screensize, yes. and [with the PDFs] you can expand them and I'm used to doing that with this [iPhone®].

Participant 2:

Yeah. When you zoomed up you could easily see it. [The charts were] moderately easy to read.

Participants felt that effective use of this app required practice and experience. The need for skill was mentioned explicitly in this participant's comment below:

Participant 1:

I was trying to get back to the dashboard and ended up going too far ... [It would be easier] once I'd learned to use it, which wouldn't take long.

Here, the need for skill was mentioned implicitly in this excerpt:

Participant 2:

I think it comes back to the user interface ... it's only as good as the person who's using and it's only as good as the information that's in

there. So, if you're rushed or busy it's easier to miss on the iPhone® only just because of the size of it. The limitations of the screen size. I'd have to take longer if I was going to use the iPhone® just to make sure I'd really read everything.

None of the participants stated that they used their own devices in medical practice.

One participant, who observed an unfamiliar chart type, i.e., the spider or stick chart used in the U.S., but not commonly used in Canada. The participant, used his or her own mobile phone to look up details of the chart; in particular, medical facts about the chart's values. This chart was shown in Figure 14.

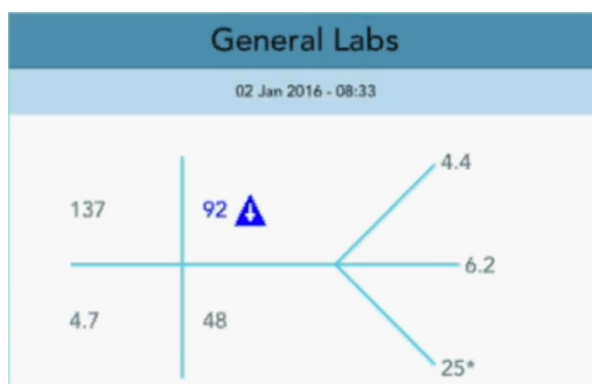


Figure 14: An example of a spider or stick chart as seen in the app

Some participants recognised this representation of laboratory results but none was able to interpret the results. One participant used his or her own smartphone sized mobile device to search the web for its meaning.

When asked about the use of a smartphone to look up medical information, the reply, that was dismissive and defensive, was interesting because it revealed both the usefulness of, and dependency on these devices as shown in the excerpt below:

Participant 1:
 Ah, but that was different, just looking up a single thing. With the app though there were many, many items on

the screen so it's too much data for the size of the screen.

Learning curve of the new system. Participants made positive comments were made regarding the learning curve (see the excerpts below). This indicated that skill was needed to learn the app, but that acquiring the skill would not be difficult. There were no negative comments on the learning curve.

Participant 1:

Um, other than the size of the font ... it's not really a problem ... just like any type of new application, it's just a matter of time to figure out where things are located but there wasn't a lot of content in it and it was simple to use.

Participant 2:

So, this little arrow feature is a good feature and I'm using the horizontal arrow that's up on vitals and when I push it goes to medications and when I look down on the four little dots on the bottom I can see that that actually moves the screen along. So that's helpful from a user perspective because you can see where you are.

Despite the limitations of screen size, participants liked the use of colour and layout of the information presented as shown below:

Participant 2:

I also think that it's a pretty good layout for that small of a device. I think having things well demarcated with different colours.

...

Right, so [regarding the symbols] it tells us what her reading was and what the normal values are. That's really

useful. The EMR we use at the hospital does not do that.

iPhone® usefulness comments. Negative comments on the usefulness of the iPhone® were coded “Impact,” defined as an unexpected problem of the system on work activities negative impact of workflow on app use (47%), followed by “Applicability”, when the information presented was not applicable to real healthcare practice or cases encountered (20%), and “Correctness” when the information presented was incorrect or inaccurate (20%).

Impact of using the device and app. One participant implied that imaging decisions made on a mobile device might be unsafe, because the ambient lighting was not consistent:

Participant 8:

I might look at imaging but I wouldn't make decisions based on imaging on one of these devices. Mostly because of the size and the ambient lighting.

The device was unlikely to have been designed for diagnostic imaging purposes. However, imagery was available for reference and discussion during patient encounters. The iPhone® app had no separate diagnostic imaging section. Images located in the reports were discoverable only when browsing through them and scanning titles for radiological terms. This caused frustration when a participant spent too much time searching for imaging, and thought the search took too long:

Participant 2:

and that there were no x-rays is a thing that we should note. I kept looking for it which took longer.

Using the iPhone® in the patient's presence was perceived as having a negative impact on workflow:

Participant 3:

I don't want to miss the opportunity to have eye-to-eye contact with the patient. I feel it interferes with the transaction and communicating effectively with patients.

One participant speculated specifically on the awareness of a tendency to make new mistakes with new technology, particularly during data entry:

Participant 4:

I'm more familiar now with the things that can go wrong when using the computer. I watch for the computer making mistakes. For example, in writing prescriptions I always check then I print out the prescription and read it and then give it to the patient.

In the comments, several participants reiterated the lack of any summary as to why the patient they were given to review was actually in hospital:

Participant 1:

But it doesn't say why, though. So how am I supposed to know...get this test from the hospital and that's all it tells me, they refuse to say, so now I have to call the patient in to see what's going on.

Participant 2:

So, the first thing I see is Vitals. What I want to see is "Patient was

admitted to hospital" you know? That's what I want to know.

Participant 2:

I still cannot tell what's wrong with her or why she's in the hospital.

Positive iPhone® comments concerned the "Portability," how easy the device is to carry (36%), "Applicability," when the information presented was not applicable to real healthcare practice or cases encountered (23%), and "Navigation," moving through the user interface (9%).

The comments on the usefulness of the iPhone running the app were negative in tone and were coupled with thoughts on how it could be useful in the context of clinical work. Several participants noted that using the app on the iPhone® in clinical work would be difficult and probably avoided. This is in direct contrast to the following section where the iPhone®'s portableness is discussed.

Applicability and portability of the app and the device. Participants liked the portability of the iPhone®. Despite comments about the iPhone®'s screen size being too small to use effectively, the portability of the device itself added greatly to its perceived usefulness:

Participant 8

Actually, I think it's better for patients to have the information at the

bedside [rather than] having to go through a wired-in computer.

Participant 4:

But so, when I want to see something specific ... that's also why I chose this

[gestures to their own small-sized mobile device]

so I can pack it and go because it was small.

Participant 2:

So, the portability of the iPhone® is interesting. The iPad® certainly is [appropriately sized]. Like 'appropriate' is everything like, hauling the thing around with you. So, there is more of a physical burden with the iPad® [and less so with] The smaller device.

Related to portability, one participant commented that the iPhone® with the app had an application outside a hospital. Being able to access results remotely was perceived as a positive aspect of the system:

Participant 5:

... if I was using an EMR where I need to put information into it, that would be time-consuming, right? But I could use it if I was at a seminar or a conference on Friday and I wouldn't want to have a laptop, I could get information off that ... critical values, I could get a sense of the person, or get an update from others, read a consultant's report

One neutral comment on the portability of the iPhone® was the suggestion to use it to look up test results while driving. This seemed a dangerous situation, but no concern was raised by the participant at the time.

Participant 1:

If I were driving along and someone phoned, I could check whether the test was right or wrong, I could use it, but it's not worth the effort as far as I'm concerned.

Navigation of the app-device. One participant commented that the iPhone®'s ability to scroll in two different directions was useful for viewing or scanning data in time series, or by categories. This was analogous to viewing a paper table where one can visually scan values horizontally or vertically:

Participant 5:

OK, so I can see that the figuring out which side to scroll so I can scroll vertically to get the different results and if I go horizontal I'm going to get the different days and that's what I'm used to in my EMR ... The horizontal view will help me if I want to compare a singular surrogate like the Oxygen Saturation over time.

The iPhone®'s received more negative comments from participants who expressed frustration while navigating and comprehending the data. The small size and inconsistent layout were frequently mentioned. The 'made-up' quality of the data elicited both humorous comments and confusion, depending on the participant's personality and their experience using the app on the iPhone®.

iPhone® error and safety comments. There was a high proportion of Neutral comments on iPhone® errors. Thirty-three percent of comments in this category were

neutral. The other two categories, Usability and Usefulness, had only 4% and 1% neutral comments respectfully.

Other comments tended to be about: workarounds (40%), mistakes (40%) and organization (20%). There were four positive comments about iPhone® errors: Workflow, Ease-of-use, Impact and Mistakes. Three negative comments about iPhone® errors concerned: Controls, Ease-of-use and Impact. Refer to Table 4: Safety and Error Coding Dictionary on page 81 for definitions of these codes.

Neutral comments. A few examples of neutral comments began with the idea of changing a value's appearance to add to its meaning:

Participant 2:

So as long as you have normal values in colour or in some way indicated and highlighted ... then the chances that something would be missed is reduced.

One participant's equated the impact of paper-use, and electronic records on public safety:

Participant 5:

Compared to using paper and people I don't know why [the app] would be any more or less safe. If someone writes down a wrong answer, or reads the wrong report or miss out this ... you're as likely to do that on paper as on [these devices]

Echoing the concern about the human side of the patient encounter, as revealed in the literature, one participant commented that although device use could affect the communication flow of a patient encounter, it was probably not a safety concern:

Participant 7:

I think the human interaction between the doctor and the patient is vital and

I think there's a tendency to interrupt it. Is that safety? Probably not ... I think it's got a place using electronic records.

When participants were challenged to address a frequently repeated Slip – inadvertently exiting the screen – they revealed their complacency and continued to use the device as illustrated by the following quote:

Participant 7:
Well, wouldn't you expect that in computers?

Usability problems abound. The foregoing comments suggest that all participants have encountered usability issues and all have learned to devise workarounds for them. Doing so is a part of electronically-assisted healthcare. All participants devised workarounds for the confusion caused by the lack of a back button, however, very few participants acknowledged explicitly using a workaround. The creation of a workaround was automatic for most participants in this study.

Workflow comments. A comment from a participant on limiting the use of a mobile device to a read-only mode was interesting. It illustrated concerns about making mistakes with medical information on a device that was seen as difficult, and at times, confusing to use (see the excerpt below). This positive comment was expressed within fifteen minutes of the session start:

Participant 5:
I guess the one thing I think about apps in general ... is like, well, apps when you actually don't—where you don't have patient information like it's not an EMR ... I don't use EMR on a hand-held, you can't screw it up, you're not going to lose information, you're just

reading. From that perspective, I never get concerned.

This idea, that read-only data on a small handheld device is safe, combined with the concept that the most useful attribute of the iPhone® is that its portability makes a more compelling use case for the iPhone® as a communications tool rather than the less compelling use case for it as a clinical tool.

Safety comments. When asked about their thoughts about patient safety while using these devices in clinical practice, another participant replied that using the device presented no difference to the safety of patients:

Participant 6:

I don't think that [the risk of being distracted] would make much difference because I'm focused anyway when I'm seeing a patient.

A different participant commented that being able to access the data readily was a benefit to patient safety:

Participant 9:

Because [apps] gather all the info in one place, quickly ... generally, I think it adds to patient safety.

A more general comment about the concept of safety of mobile devices came from one participant who suggested the potential for greater improvements in patient safety could be afforded through the use of mobile devices. One participant referred to the potential for a decision-support system within an app:

Participant 1:

Thing about [mobile devices] though is that you've got the ability to build in [validation]. So, you could say "so this is a value outside the limits, did you mean it", or "the last two patients

we gave this to died", or something like that. So, your ability to build in fail safes is higher with the devices than it is on paper.

Also, implicit in this comment was the use of natural language to describe the effects of medication errors. After the session, one participant, who considered using a mobile device for completing post-encounter paperwork, became concerned regarding the size of the Controls and Fonts, and expressed anxiety about losing data by using the iPhone®:

Participant 5:

The one thing I'm concerned about in the future using an app to actually do my documentation -- making sure I don't accidentally flip a little [control] with a light touch and lose something.

The need for experience, training and perseverance, especially in relation to interoperability, is seen as an impediment to the safe use of mobile devices as described by participant 9 below:

Participant 9:

but because [mobile devices] are technology, they have some difficulties working with other technologies, you need to be familiar with the options and the functions, and you need to be patient and not get frustrated and put it away.

Participants indicated that the iPhone version of the app was more difficult to use and would take time to learn its controls and navigation. Much of the frustration appeared to be caused by both: (1) the quality of the patient data within the training mode of the app, and, (2) the unfamiliar navigation controls provided by the app.

The iPad®

This section of the chapter describes the results from the participants' interaction with the same app using the iPad® shown in Figure 16 below. The task given to the participant: "Please review the data and summarise the patient's health" was the same as that given for the iPhone,® except that they were assigned a different patient to review.

On reviewing and coding the audio and video evidence, there were a total of 238 codes recorded. These were, in descending order: Usability issues (67%), Usefulness issues (26%), and Safety issues (7%). A summary breakdown of these issues by category and code-type is shown in Figure 15 below. A full breakdown is also shown in Table 6 below. The greatest Usability count was Comprehension and Consistency; the greatest Usefulness count was Correctness; and, the greatest Safety count was Slips. Refer to Table 2: Usability Coding Dictionary on page 80 for definitions of the foregoing codes.

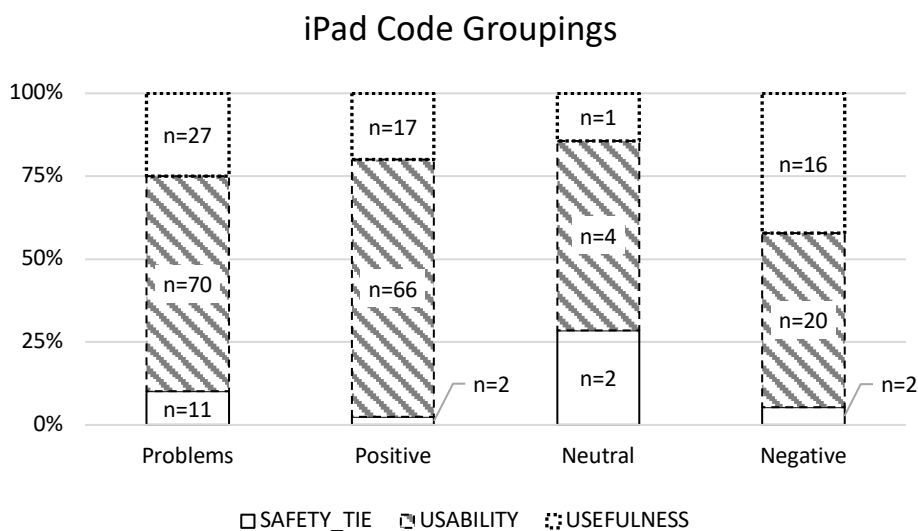


Figure 15: Percentage distribution of iPad® codes by category and type

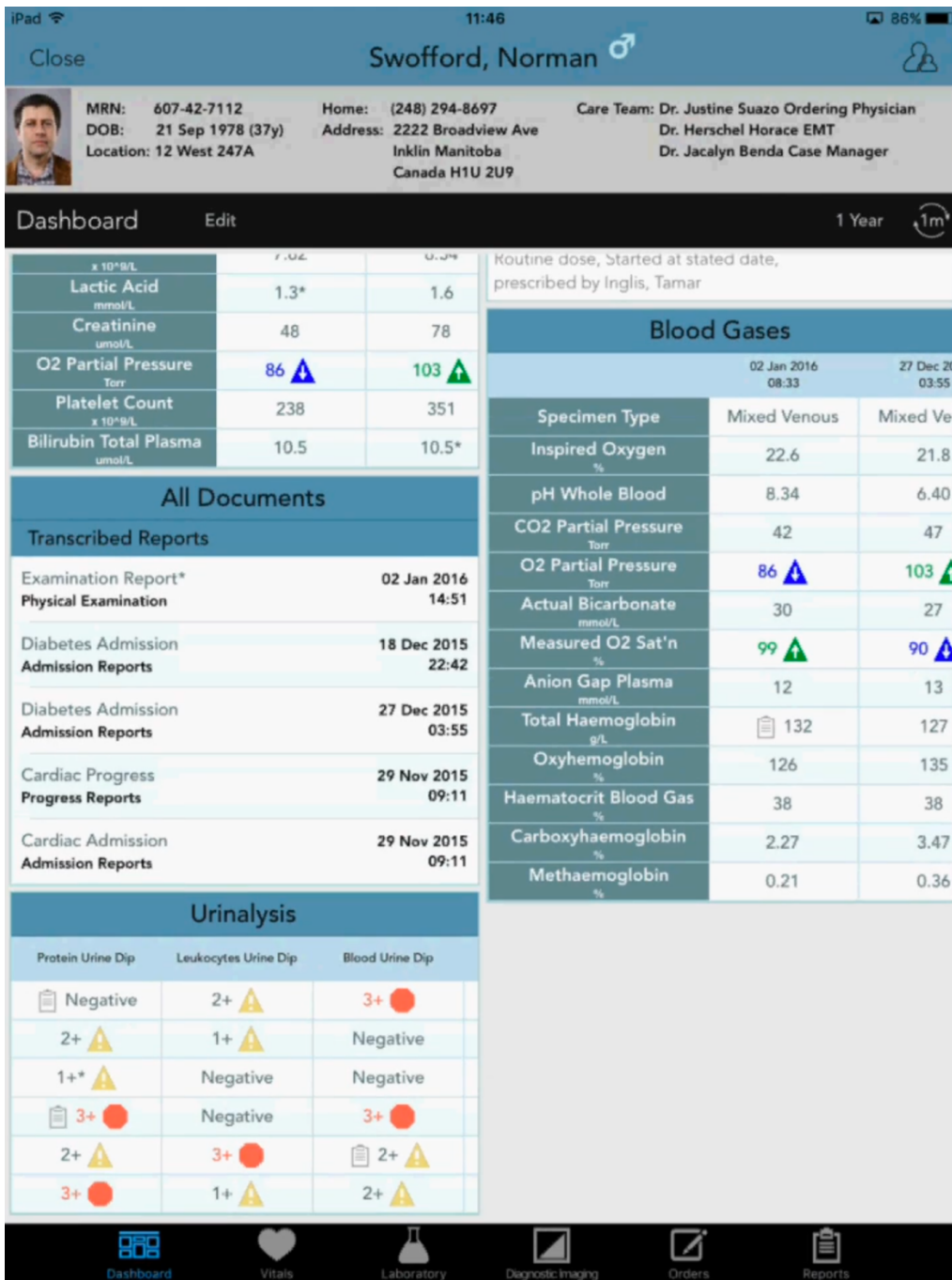


Figure 16: The iPad® app showing the dashboard

The dashboard is the view that participants arrived at after selecting a patient from the list. The main navigation control is the persistent row of icons along the bottom. Clicking these navigates the user interface.

Table 6: Problem and comment scores by code for the iPad®

<i>Codes</i>	<i>Problems</i>	<i>Positive Comments</i>	<i>Neutral Comments</i>	<i>Negative Comments</i>
<i>Error</i>				
Mistake	2	1	2	1
Navigation	2	0	0	0
Slip	7	0	0	0
	<i>11</i>	<i>1</i>	<i>2</i>	<i>1</i>
<i>Usability</i>				
Colour	2	1	0	0
Comprehension	13	0	0	2
Consistency	13	0	0	1
Controls	1	2	0	1
Ease of Use	1	16	1	0
Familiarity	2	5	0	1
Font	0	6	0	1
Graphics	5	3	0	3
Instructions	1	0	0	0
Layout	5	7	0	0
Learning Curve	0	0	1	0
Navigation	12	16	0	1
Organization	4	2	1	0
Screensize	0	5	1	6
Terminology	8	0	0	1
Workflow	2	2	0	3
	<i>69</i>	<i>65</i>	<i>4</i>	<i>20</i>
<i>Usefulness</i>				
Applicability	4	5	1	4
Correctness	14	2	0	0
Impact	2	3	0	3
Portability	0	1	0	6
Relevance	4	3	0	1
Timeliness	1	0	0	2
	<i>25</i>	<i>14</i>	<i>1</i>	<i>16</i>
<i>Total</i>	<i>105</i>	<i>80</i>	<i>7</i>	<i>37</i>

Participants encountered a total of 105 problems on the iPad®. The majority of these (n=69), were categorized as Usability issues, followed by Usefulness (n=25) and finally

Safety issues (n=11). Refer to Table 2: Usability Coding Dictionary 3, 4 and 5 starting on page 80 for definitions of the codes listed in the table above.

iPad® usability problems. Of the iPad®'s Usability problems, 19% were concerned with Comprehension, 19% with Consistency, 17% Navigation, 8% Terminology, and the remaining 45%, assigned to other codes, see Figure 17 below. Refer to Table 2: Usability Coding Dictionary on page 80 for code definitions.

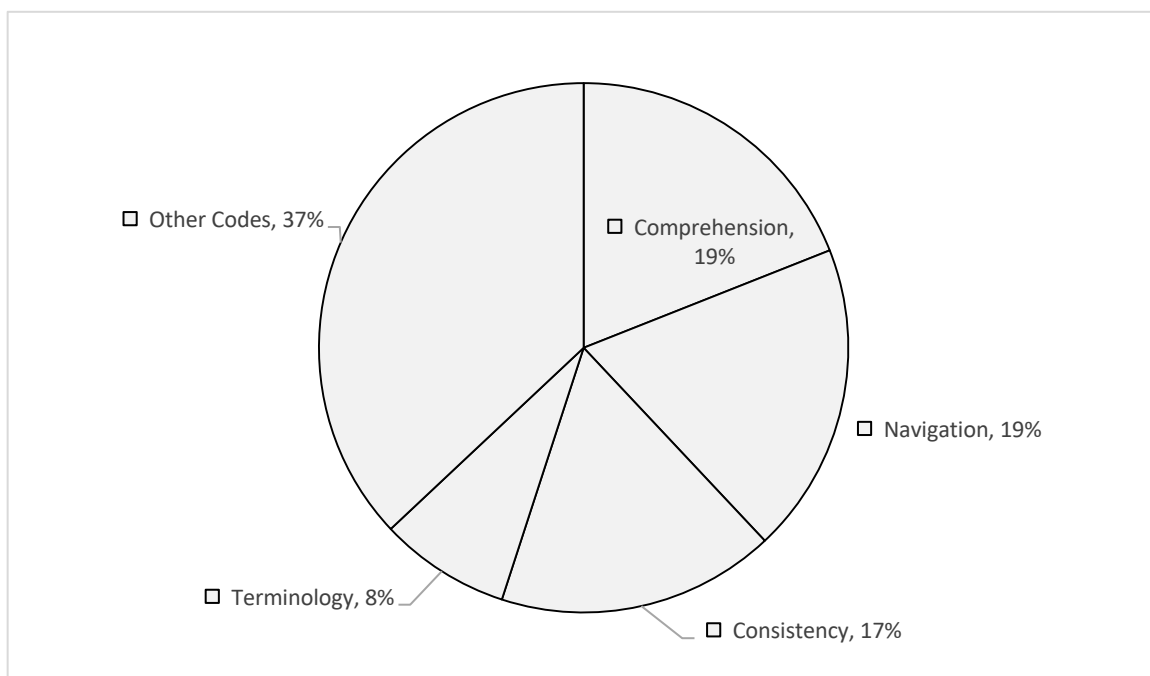


Figure 17: Percentage distribution of iPad® Usability codes by category

Comprehension. Nineteen percent of participants noted they had difficulty understanding icons. Lab values that were shown to participants had a number of text decorations applied to give emphasis to, or add additional meaning to the user interface design. Following convention participants understood that a red warning symbol generally meant ‘one should pay attention’ {Anonymous:2012vd}. A number of other

symbols were shown to participants. No legend was provided to inform the participant as to their specific meaning. These symbolic text decorations caused a confusion for some participants as the following excerpt illustrates:

Participant 4:

They've put an asterisk around it and I don't know what the asterisk means ... and ten point five is asterisked and the [same value] is not asterisked on the 27th. Don't know what that means. I would say that this guy is thirty-eight and he has cardiac and diabetic problems. He has problems and he's not looking after himself very well. Blood urine drip. It has red exclamation marks that I suppose we're to be looking at. He's got a higher than normal .. well the red stop sign generally means whoops. and I don't know what the *Negative* of course it isn't really negative so why does it show a chart next to that and not the other ... Too many poor symbols signifying nothing as far as I'm concerned.

The confusion was exacerbated by the fact that the app used U.S. customary units. As a result, the unclear text decorations, coupled with unfamiliar units, led to this participant's inability to comprehend the state of the patient's health:

Participant 8:

So, he is in fact diabetic... checks his sugars in American numbers and I don't know what they mean but they don't look too bad.

In addition to icons and colourful text decorations (see Figure 18 below), the designers of the app also used flags, words styled as headings, to indicate the status of the lab results page. Since the flags had a similar appearance to clickable or 'tappable' elements, confusion occurred when users encountered flags on lab results:

Participant 8:

I don't know what that is.
 "Discontinued" what does that mean?
 They asked for that and it wasn't done?
 I don't know what they mean by that. [I
 wonder if it means that] they did they
 it once and then stopped doing it.

Consistency. Nineteen percent of participants identified consistency as an issue. Most data tables indicated the date of the value; some table rows did not indicate dates. Urinalysis tables, shown in Figure 10, had the greatest number of icons and decorated text, but no dates:

Participant 2:

There was a place in the Urinalysis section that makes it unclear, like there is a column or row missing like the dates so you cannot tell when the data was taken.

Participant 7:

Oh, there's protein in his urine. Though there's no indication in the urinalysis piece because there's no date when they were taken. So that's a little weird.

Participant 8:

and the orders ... there was a perfusion test ... [it does not] tell

me when these were ordered. This one was discontinued. ...

Urinalysis		
Protein Urine Dip	Leukocytes Urine Dip	Blood Urine Dip
Negative	2+	3+
2+	1+	Negative
1+*	Negative	Negative
3+	Negative	3+
2+	3+	2+
3+	1+	2+

Figure 18: The Urinalysis table as seen on the iPad®

The Urinalysis table shows lab results in a heavily decorated iconographic style. There are plain-text, coloured text, plain numeric values, and coloured numeric values. Additionally, present in the table are white clipboard icons, yellow triangles with embedded exclamations, red octagons, plus symbols and asterisks. This symbol-dense layout is scrollable left and right to reveal more columns. No legend is supplied to aid the user's comprehension.

The app used combinations of colours and icons, i.e., a yellow triangle indicates a warning and a red octagonal symbol indicates a likely more urgent warning. Some participants became frustrated with the large number of warning symbols. This observation produced a constructive comment by a participant regarding the usability of icons and their ability to indicate an abnormal value. This participant suggested using a single symbol and altering its colour to encode severity:

Participant 9:
Yes, I can just click on them.

[taps a table cell with a red icon]

ah 'critical' yeah so clicking on them works. and if I click on [a cell with a yellow triangle] then it's 'abnormal'

so there's a degree of abnormality that goes from abnormal to critical

If I was going to grade something in the same category then I would use the same symbol with different colours. I wouldn't change the symbols, just the colour.

An inconsistency existed between the iPad® and iPhone® app. Although the app on both devices had imaging reports with embedded imagery, the iPad® alone had a separate section for the imaging itself, i.e., without the reports. One participant noticed this feature:

So those are the images but not the reports...just this button here.

This particular inconsistency caused most participants to miss one, or the other, section when using the iPad®; particularly, when they encountered the imaging section, before the reports section. This pattern led some to assume there were no imaging reports as illustrated in the quote below:

Participant 3:

With some more practice with them yes, I think I would. Probably two days would be enough. I see about forty people a day. The fact that there isn't a report about the x-rays is something that sticks in my mind. They show the images but not the reports. That's not very comprehensive.

One participant noticed that the documents in the Reports section of the app had formatted PDF reports, that seemed to be missing page numbers. This was the case only with single-page documents. All multi-page documents displayed both page number and total page count, i.e., "Page 2 of 3". In the case of the single-page document; however,

no page number was displayed. As a result, this participant was uncertain about the number of pages in a document, and wanted to see an indication of the total page count, i.e., “page 1 of 1,” as Participant 8 describes below:

Participant 8:
and the orders ... there was a
perfusion test ... tell me when these
were ordered. This one was
discontinued. ...

[navigates to a different report]

The '1 of 2' was very useful but I
would like to see a '1 of 1' here.

Navigation. Most iPad® app navigation problems (n=9, 17% of participants) were caused by the lack of a familiar navigation control: the ‘back’ button. The iPad® app lacked a ‘back’ navigation control that allowed the user to return to the previous view; instead, the user was required to navigate to a specific view. To do this, the iPad® provided a navigation bar on the bottom of the screen. This feature did not provide the familiar ‘back’ navigation control. Some participants did not use the navigation bar, shown in Figure 16. Instead, they scanned the user interface for a ‘back’ button; and became frustrated when none was found. Some participants assumed that a ‘close’ button

had the same functionality as a ‘back’ button. When participants tapped the ‘close’ button, the patient view unexpectedly closed, as illustrated in the following excerpts:

Participant 9:

I wanna go back so I ... refresh? No.
Close? Oh! I've gone out [and returned
to the patient list]

Participant 8:

That would be my suspicion ... oops

Participant 7:

oops now I've messed it up

Participant 2:

... and how do you get back?

Participants were jarred into re-establishing their location within the app and forced to restart their navigation to return to a recognisable UI. Most participants quickly gave up solving their location and resorted to starting over from the beginning by using the patient chooser to re-select their patient. No participant chose the wrong patient when presented with the list of patients.

Terminology. The iPad® app was designed for use by in-hospital physicians; however, not all participants worked in a hospital setting. The app's fundamental function was to connect to an EMR and retrieve data. Some participants were specialists, Some were general practitioners. Eight percent of participants did not understand the app's terminology. As a result, not all terminology was immediately clear to all participants as seen in the excerpt below:

Participant 6:

... No abnormalities. Except for an
exaggerated ... no it's not he's in

flexion. A bit of osteoarthritis. What is this Acid perfusion test?

The app provided no way for the participant to look up the meaning of a term within the app itself. For example, one user did not recognize the name of a medication found in the patient record as seen below:

Participant 7:
 "Pamidronate" ... no idea what that is.

The Urinalysis table contained no legend to explain the meaning of its icons and colours. No explicit terminology or legend existed for this table which frustrated one participant:

Participant 6:
 Urinalysis ... [this patient] has ... are these [data values] for different dates? There's 3+1 there and I don't know what that means, there's six different urinalysis values here.

The use of U.S. terminology was also an issue. The spider chart, shown in Figure 14, used in the U.S., but not commonly used in Canada, appeared in the iPad®. Some participants recognized the chart, but none knew how to read it, as described below:

Participant 4:
 General labs, the document here is basically a cross with an arrow on the right. and it's not clear to me where this comes from. There's [a value within it] that's highlighted with a down-arrow within it. Which might ...

Some screens used the United States customary system of units. One participant commented that it was unusual to see this, and expressed a general frustration:

Participant 1:
 The units, they're all in pounds, not kilos, six foot two, etc. The [U.S., is] only country in the world that uses

these units which makes it difficult at international meetings.

The presence of U.S. terminology and units was surprising to some participants who noticed the change after viewing the patient's Canadian demographics (as described in the following section). Some saw the Canadian addresses after the U.S. units. The manufacturer of the app is a Canadian company located in Toronto.

iPad® usefulness problems. Approximately half the iPad® Usefulness problems participants encountered concerned were coded as Correctness, i.e., the information presented was incorrect or inaccurate. Data errors caused distraction. When encountered by participants, incorrect information caused the participants to pause and comment on the fact before continuing. Refer to Table 3: Usefulness Coding Dictionary on page 80 for the meanings of codes used in this section.

Correctness. Fifty-two percent of participants identified correctness as a problem. Patient medical "Reports," i.e., physician-created documents available for review, in the app appeared to be from Canadian sources, i.e., patient demographics indicated Canadian

addresses. Some details, however, gave the impression that they were incorrect or fabricated such as the incorrect telephone area code seen in the following excerpt:

Participant 4:
Ok so I have a 37-year-old patient
who's almost thirty-eight. Is 2-4-8 a
Manitoba phone number? ³

While looking for normal limits of a lab test, one participant discovered several limits with highs and lows of zero as outlined by the following quotation:

Participant 9:
Here is one. Normal limits zero to
zero. Maybe there are too many zeroes
here.


Surprising values for lab measurements were deemed by the following participant to be impossible for a living person:

Participant 8:
Ok. So .. looking at his date of birth
he's 37 fairly young gentleman ...
looking at his vitals ... he's not
tachycardic ... he's ... oh! He's got
an oxygen saturation of a hundred and
four ... which is impossible ...

One participant observed that a report for a shoulder surgery contained incorrect imagery of a hip x-ray. This was illustrated with a screenshot (see Figure 19) of this participant's session and with the following interview excerpt:

³ Manitoba's area codes are '204' and '431'

Participant 7:
 Plain radiographs of the right
 shoulder. Post-op rotator cuff repair.
 [image shown is clearly not a shoulder]
 I think they must have shoved in a few
 images just for fun.



Good Health Hospital

30 University Ave.
 Toronto ON
 Canada
 M0H 1Z5
 1-800-555-9533

Musculoskeletal X-Ray

Admitting Physician: Dr. John Granger
Attending Physician: Dr. Marilyn Stanhope
Family Physician: Dr. Frank Thomas

Document Id: 5-52344-TZ
Order Number: 95505553
Referral Code: A37-531

PLAIN RADIOGRAPHS OF THE RIGHT SHOULDER

CLINICAL HISTORY: Post-op rotator cuff repair.

COMPARISON STUDY

FINDINGS: Interval rotator cuff repair with three soft tissue anchors and on the anterior aspect of the humeral head. Skin staples and soft tissues swelling in keeping with recent postoperative changes.

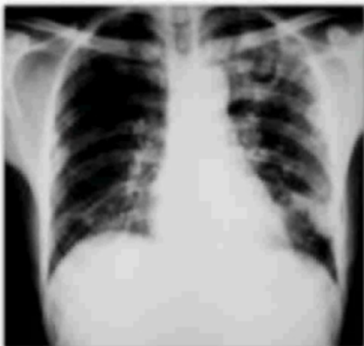


Figure 19: Imaging report with an incorrect X-ray image

A screenshot showing what a participant saw when choosing to view a patient's shoulder X-ray report. Notice that this report described a shoulder imaging study but featured a chest x-ray image that did not show either of the patient's shoulders. Note also that the patient's name was not shown.

Three participants commented on reports labelled as admission, outpatient, and inpatient. On reading these documents, some participants identified that the content of the documents was actually outpatient mislabelled reports. It became clear to the participants that the labels were not accurate and not to be trusted as illustrated in the quotes below:

Participant 7:

Ok here's one. ... Diabetes
Hypercholesterolemia and sinusitis ...
so this is not [an] in-hospital
[report] it's an ambulatory visit

Participant 2:

Now the admission report looks like an outpatient visit it's not an admission report for acute care. But it does look like from the other information that he was admitted for acute care. They call it an admission history but I'd say it was more like a clinical note

Some of these errors led to comical comments, and contributed little to the perception of the app's trustworthiness. For example, two participants noticed that a one patient's report contained images from two different individuals:

Participant 8:

So he's has a chest xray. and I don't see anything acute going on ... but .. oh what he's got a pacemaker... on the lateral but not on the AP! So I

would question whether that was the same patient.

Participant 1:

There's his problem... He has that nice wire in his heart. This x-ray likely doesn't belong to this patient.

When one participant was asked by the researcher: "Is that typical for a cardiac report?" the participant replied as follows:

Participant 1:

No! Unless they were written by a dietician.

Subtle locale errors, inconsistent use of units of measure, inconsistent systems of units and missing normal ranges for lab values caused some participants to question the vitality of the patients they were assigned to review. The realisation that participants could not tell whether the patient was alive or dead caused some participants to take a break from the gravity of the chart review and make jokes, implicitly suggesting that the app's data was not to be taken seriously.

Relevance. Fifteen percent of problems were coded as Relevance issues, i.e., where the information was perceived as being irrelevant or inaccurate. Some issues related directly to the issue of the app having been designed to be used in a hospital. In one case, the information presented was of little use for a doctor who was not a specialist as outlined in the quote below:

Participant 3:

For me and my type of practice [GP, walk-in clinic], the angiogram and the CT lung don't provide me with any useful information. Because I'm not

used to viewing those types of images.
...

So, the arterial is the one I'm interested in and that was 86, lower than I would have expected. Here are umm for the mixed venous which doesn't mean anything for me.

The app contained data with many values for many lab tests as well as monitoring data. Some of these values, i.e., weight, were important to have, but was seen by some participants to be of limited use. According to one participant, a patient's weight was often used not solely but combined with other data, to derive a third value, for example, the patient's Body Mass Index (BMI). The BMI is calculated from the patient's mass and height and was expected to be shown alongside the patient's weight.

One participant commented that there was an issue with the efficiency of using the app as a whole. They pointed out that repeated navigation between sections of the app to build up an impression of the patient's health, was not efficient because of the high number of irrelevant items they had to navigate in order to form an impression as described by one participant below:

Participant 8:

Let's go back and do another one.
You see this is very annoying I find:
Past medical history: refer to chart;
medications: refer to chart; allergies:
refer to chart. What does that mean?
That means time. Inefficient, in my
opinion.

iPad® error and safety problems. On the iPad®, most of the errors made by participants were Slips (64%). Slips are defined as errors, recognized and corrected by the participant. Mistakes are errors that were not recognized by the participant, were

18% of Error and Safety coded problems. Refer to Table 4: Safety and Error Coding Dictionary on page 81 for the meaning of the codes used in this section.

Slips. One participant misinterpreted the chronological order of the results presented by the app, that led participants to think that the patient's condition was deteriorating; when, in fact, the patient was improving as illustrated in the quote below:

Participant 9:

I see this patient was on humidity ... room air and [then] needed to be put on a tracheal mask and then Venturi Mask and finally on a ventilator .. so maybe the patient became more critical. Over [these days] ... Oh it goes backwards! I don't know why but I expected it to be the opposite.

In another case, data present on the screen, was missed, then found. One participant read the patient information on the app. This patient had a list of ailments which suggested he or she was an older patient. Participant seven expressed concern and desired to confirm the patient's reported age. Although this information was actually shown on the currently displayed page, this participant missed it and made the decision to navigate to another page to retrieve the patient's age as seen in the excerpt below:

Participant 7:

So here: neck pain radiates to the right extremity. Numbness ... motion diminishing re ... well what age is this person!?

[misses patient's age displayed on screen. Navigates and finds the patient's age on main page]

She's 21 years old ... that's odd I thought she was young and how come

she's got a degenerative disease in the neck and so on.

The remaining Slips had to do with navigation issues, whereby the participants tried to go 'back' to a previous screen, and inadvertently quit the patient view returning to the patient list. As discussed previously, there was no 'back' button on the iPad® app. This was flagged as a Slip because it was an unintended action. All participants who made this Slip, took corrective action. Some used it intentionally as a Workaround.

There were slips made, most likely as a result of a lack of experience with app. Most of these slips were caused by participants who navigated unexpectedly when they did not intend to, or by navigating to a different part of the app when there was no reason to do so. In both cases this mis-navigation had a detrimental effect on the participants' efficient use of the app.

Mistakes. The code Mistakes referred to errors not recognized by the participant. Mistakes appeared as a code 18% of the time. The researcher explained to one participant that the meaning of the status icons on a particular table indicated the value was abnormal. The response was:

Participant 7:
Ok, so I assume that everything else is normal because they're within the normal limits.

This participant assumed that values without icons were normal values. Icons were not used in every data view. Some tables with values did not use icons, even if the values were out of range. This participant's assumption that the lack of iconography represented that all values displayed were normal was false. The lack of icons did not indicate the lack of a problem and this resulted in an error of interpretation by the participant.

Another instance of missed information occurred because an allergy list appeared without icons. One participant misinterpreted this as a list without checkmarks, i.e., no checkmarks, no allergies; whereas, the list without the icons, was actually the list of the patient's allergies as described by the participant below and illustrated in Figure 20 below:

Participant 7:

OK 37 year old patient, so he was admitted on the 18th and [starts reading data] blood pressure looks fine, on the 27th he had a slight fever ... but on the 22nd that has decreased, heart rate has increased, oxygen saturation is quite low initially but then it got better so I guess he was given some oxygen, **allergies are none.**

Misinterpreted data had two main causes: (1) the participant added meaning when none was present for example no icons equates to no problems, and (2) that a lack of icons invalidated a list of items. Both misinterpretations represent a problem of consistent use of visual aids in the GUI and both represent a pathway to potentially serious error.

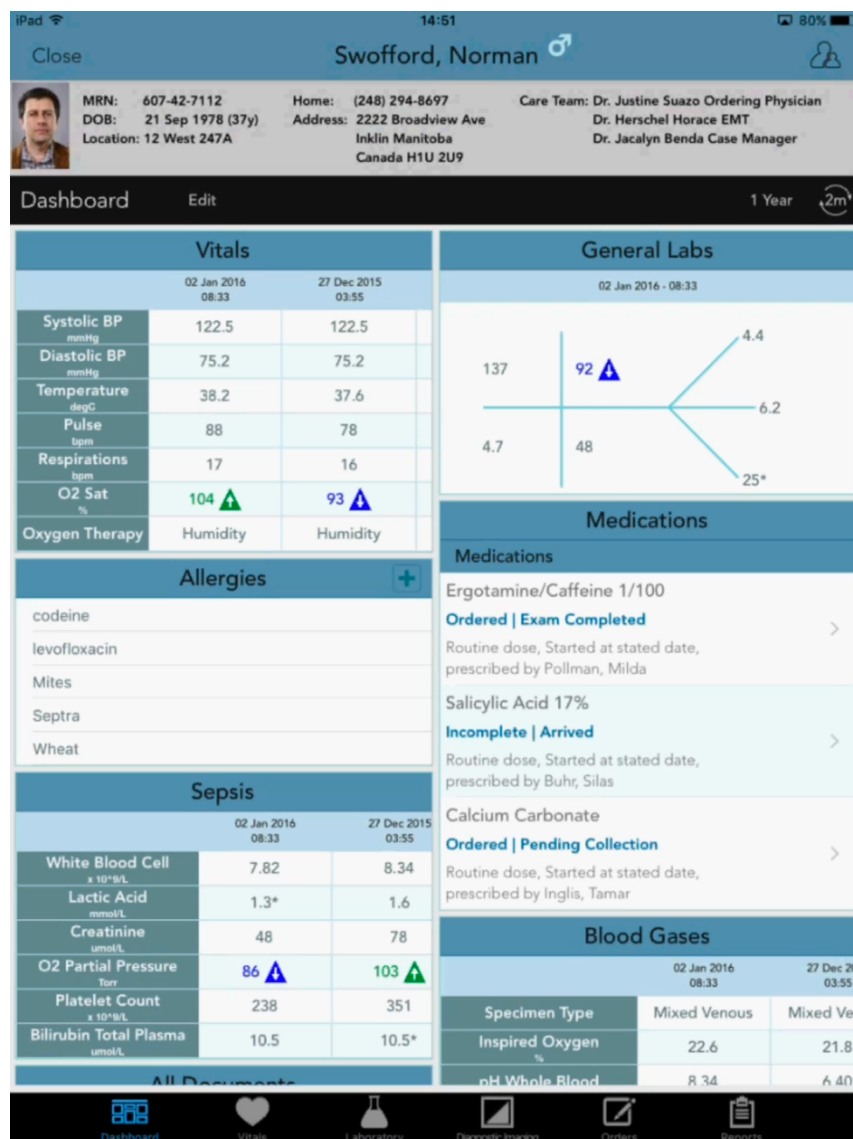


Figure 20: Page view showing a list of allergies centre-left

One participant interpreted the list of allergies show as "Allergies are none". The undecorated list had conveyed the concept: none on the list applied to the patient.

iPad® usability comments. Comments were made by the participants during the interview part of the sessions. In this section, the results of comments made throughout the session will be presented. On the iPad®, there were 130 comments: 69% Usability, 26% Usefulness, and 5% Safety. Of the Usability comments, 19% were coded as Ease-

of-Use, 19% were coded as Navigation, and 13% as Screensize. Refer to Table 2: Usability Coding Dictionary on page 80 for definitions of the code terms.

Ease-of-use. Participants found the iPad® app intuitive and easy to use. Information was available and accessed by tapping around the user interface. One participant stated that this was the case since most users would already have the skill to use the device. They implied that the learning curve for the user would be lessened by not having to become familiar with the iPad® because most physicians have experience with such devices. On the other hand, being unfamiliar with a device could have an effect on new users. See below for statements made by participants:

Participant 2:

Yeah, [using the iPad® is] pretty straight-forward. If you didn't know anything about iPad®s it'd be a nuisance. But there's not many docs that haven't used them. Yes I would say I did. Well I find there's a good balance of information on the screen

Participant 1:

If I was looking for "has this patient changed over time" this [iPhone®] is not as good as the [iPad®] because you have to scroll a lot to find them. Because you can't see more than four or five in a row.

Despite the positively perceived ease-of-use of the app, the app itself did not provide a concise acute-care summary as seen in the quote below. One participant commented that despite the easily accessible historical data, nothing was found that informed them of the immediate reason the patient was in hospital:

Participant 2:

It was relatively easy to use it was relatively intuitive. I could find most of what I wanted to by tapping all around. I could find a lot of the historical data. What I wanted was the acute, the right-now and that's what I couldn't find.

One participant implied that not all potential users of the iPad® or computers, would be sufficiently computer-literate to use them successfully. Some of these users, especially those who work in small clinics, i.e., physicians, would benefit from IT support:

Participant 8:

There's a lot of things in the computing world that if you're a GP in an office on you own and it could be very annoying. You'd do something and you'd be stuck and there's no Siri to ask "what's the goddam asterisk mean?!"

Participants comments indicate that the app was perceived to be easy to use but some participants acknowledged that skill was needed in two forms: (1) general computer literacy skills and (2) app-specific skill. The latter would take time given the complexity of the app's user interface.

Navigation. The navigation bar on the iPad® was always shown to users, who needed to tap only once to switch to a different view. This fact was commented on by one participant when asked about navigation on the iPad®:

Participant 8:

The navigation ... that bar was on the iPad® all the time on the screen.

The navigation bar on the iPad® app, rather than being hidden in a modal popup dialog box, was always present in the display, as related in the quote below:

Participant 4:

Yeah but once you've done that a couple of times you remember. So I mean doing it the first time. I mean this is an intuitive thing ... the 'Apple' is intuitive ... well some people have different intuitions than others.

The larger display of the iPad® had more area in which to place tables, charts and lists. Users had less cause to interact with the app and view multiple pages around; as a consequence, they had fewer reasons to navigate as described by one participant below:

Participant 7:

Hmm. I don't about problems, but the iPad® never seems quite as touchy as the phone, I don't know why I think that, but maybe it's just my brain that thinks because it's bigger. But doesn't see quite as touchy.

This participant stated that there was less need for manual interaction, and fewer reasons to fiddle with the UI in order to assimilate the patient's information.

Screen size. One participant commented that they preferred the larger screen. It made the app easier to use because it required fewer navigation actions to collect all the information they needed as described below:

Participant 1:

Oh, it was much easier. and you get more information on one screen on different sections, not flipping from section to section. I have information in front of my face. Everything I need is right there. All I have to do is

click, click, click to wherever it is.
and there it is. It's way superior.

The sections shown on the iPad® screen were large enough to contain more rows of patient data. One participant's comments indicated that they had less scrolling to perform in order to see the values, see the comment below:

Participant 7:

I would say the bigger the better in practice. On a larger screen you can scan the information presented, on the smaller one you have to scroll a lot.

The greater size of the iPad® was still not sufficient for one participant, who was used to performing clinical exams while using a full-sized computer monitor. The participant was aware that there was an even larger iPad® on the market. This participant commented on the iPad® not being large enough, and expressed a desire to use the app on the larger device. The comment contains an implicit warning that insufficient information can lead to patient harm and an explicit consequence that such harm could lead to death as described in the excerpt below:

Participant 1:

I'd like the super-tablet! You could get more information on there. The [other] iPad® with the larger screen. The amount of information you're getting on the iPad® with this app is minor. I wouldn't make a diagnosis based on that. It's just not enough. I'd need more [information] before prescribing drugs. I could kill you just as soon as I could cure you.

iPad® usefulness comments. The iPad® Usefulness comments were grouped as follows: 29% Applicability, 21% Portability, and 18% Impact. Refer to Table 3: Usefulness Coding Dictionary on page 80 for the code definitions.

Applicability. One participant commented about the lack of derived values next to patient data. Derived values are calculated from existing data to provide more meaning for the physician to use when evaluating a patient chart. One participant stated a desire to view the anion gap, derived from patient lab data, and different participant, looked for the patient's Body Mass Index. Both stated a desire for derived data that was not available within the screen they were viewing as described in the following quote:

Participant 7:
185 pounds, so she's a bit overweight,
the thing I look for [on a report] is
the body mass index

Many of the fields in the rows of patient data contained metadata, such as reference ranges for laboratory results. The presence of metadata was seen by this participant as beneficial and applicable to their knowledge of the patient:

Participant 5:
I was pushing this button because I
wanted to see if it was true those were
the only three medicines he was on.

One participant commented that the dashboard was applicable to the task of navigation through the app's data:

Participant 9:
 So this dashboard has seven sections
 ... so I can navigate and see
 everything.

Some user interface components were perceived to be applicable to some participants. The derived data was mentioned more than once and could be read as a trend, although extrapolating, each potential user could wish for a unique particular calculated data result. This could lead to an over-abundance of information should all such requests be honoured.

Portability. Portability was discussed by many participants. The standard iPad® weighs less than half a kilogram and is 24cm by 17cm and is 7.5cm thick. The device was smaller than, and about as heavy as, a typical clipboard-style chart holder. Comments on the size of the iPad® indicated that participants liked and preferred the size of the device; but they also expressed the opinion that it was not sufficiently portable; for this reason, they consistently preferred the iPhone® device as shown below:

Participant 2:
 Hauling the thing around with you ...
 So, there is more of a physical burden
 with the iPad® than the smaller device.

Participant 6:
 I Find the iPad® more useful when using
 it but I don't like carrying it about.
 I like the smaller size. No one's going
 to carry the iPad® around.

One participant had a specific workflow for making notes in the patient's chart. This workflow made use of a speech-to-text system installed on computers in each exam room of the clinic. The software was not available on the iPad®s and therefore negated the perceived usefulness of the iPad®. Had it not been for that limitation, the iPad® would

have been preferred for the size of the display, not for its portability as described below by one participant:

Participant 4:

If I was wandering about in the ward checking up on things then I would prefer the small smartphone app because it's so much easier to carry. But if I was taking a detailed history and I couldn't use DragonSpeak then I certainly would find the iPad® easier because it's bigger and you can see things more easily and you can put a whole [document] up if you need to. So in my own practice? I'd run a server and run both off of that. But for wandering about, I'm not a carry-an-iPad®-with-me kind of [physician].

Not all participants negatively perceived the iPad®'s portability. One participant commented that the portability of the iPad® was beneficial to patients who were not ambulatory as described below:

Participant 8:

Actually, I think it's better for patients to have the information at the bedside [rather than] having to go through a wired-in computer.

The iPad® is perceived as being less portable, even cumbersome to carry when away from the patient or clinic, however, it is also perceived as being better suited to viewing patient data. The iPhone, in contrast, is perceived to be significantly more portable but not as useful for reviewing patient data due the large amount of data and the limited screen size.

Impact. One participant commented on the iPad® and suggested that the device could have an impact on their workflow, by allowing data entry and charting to be performed locally to the patient, see below.

Participant 3

The iPad® would let me do both [data-entry and reading documentation] more easily, likely could do some charting quite a bit easier.

Another comment suggested that the impact of the iPad® on a patient encounter would be minimal, and added that the encounter might improve as a result of using the iPad®'s content as a visual aid:

Participant 4: Given that people use laptops or even monitors ... I mean, will the iPad® get in the way of you talking to your patient? Well, no more than the paper you were writing on in the first place. It might be better if you are having to explain things to a patient, the iPad® might be a better thing to do it with.

This was a somewhat abstract notion, and granted equivalence to whatever the physician used to contain patient-related information, i.e., a clipboard with a chart, a set of paper notes, an iPad®—all being an object used for the patient by the physician.

This participant also commented on the negative impact the iPad® could have on body language:

Participant 3:

I don't want to miss the opportunity to have eye-to-eye contact with the patient. I feel it interferes with the

transaction and communicating effectively with patients.

One participant who worked with geriatric patients had a negative view about using the iPad® with older patients. The concern related to the potential confusion some older patients might experience with the new device. The participant advocated implicitly for a more technologically conservative approach, i.e., paper, in encounters with older patients – because ‘they’re just not ready yet,’ as seen below:

Participant 5:

Well, it might be, [good with] the older population, people who are 55 years old and older, 90 year olds ... I could totally see using these as the population ages ... all the different technology, but if I were working with younger people they would have this stuff on their own devices and so why wouldn't you use technology based on the patient population?

Conversely, the following participant thought that their existing technology, a computer in the exam room, caused problems with safety and confidentiality during patient encounters because the computer display was visible to the patient. This scenario was perceived by the participant as a positive impact on both their workflow and the patient's confidentiality (see below):

Participant 5:

[In clinic,] my computer's on. and there's always issues with that because you're doing your best to have it on their chart and do your best to shut it down, and the patient's talking to you but these devices might make better because it's not easy for the patient to see as a desktop. In some ways it

might be better for patient safety and patient confidentiality

The usefulness of accessing imaging on the iPad® was perceived to be more efficient than what was offered by their current workflow, as outlined by a participant (see below):

Participant 2:

That's pretty good. I'm no radiologist but there doesn't seem to be any pneumonia in there. That's actually a pretty nice feature there. Our EMR doesn't easily show us imaging.

One participant recalled that during training as a physician, they were required to commit a great many facts to memory. On using the reference ranges shown in the lab results section of the app, this participant had the insight that some of this memorization is redundant now that a device can provide the information at a tap of a button as seen in the following excerpt:

Participant 9:

I come from an age in which we had to memorize most of this information, so having something this size ... is a miracle ... and it's worthwhile. Yes, we needed to [this information] them, but we didn't memorize them [laughter]

One participant commented that as a result of experiencing many clerical errors, they learned to double-check that results were as expected as shown in the following excerpt:

Participant 3:

I'm more familiar now with the things that can go wrong when using the computer. I watch for the computer making mistakes. For example in writing prescriptions I always check then I

print out the prescription and read it
and then give it to the patient.

A trend in comments, that suggests participants perceive older patients are less capable with technology. Several participants were members of an ‘older generation.’ Despite participants’ perceptions of older individuals’ technological capabilities, paradoxically, there was no correlation between age-group of the participants and their facility with the devices. In this research, the oldest participant had the greatest ease with the hardware and app despite not being a regular user of neither the iPad nor iPhone, and, had never used the app before.

iPad® error and safety comments. There were five comments in the Safety category. One participant’s medical practice catered to inner-city marginalized adults. The participant’s response to the question: “What are your thoughts about patient safety?” was to consider that an expensive, easily-stolen device, reduced the clinician’s physical safety since it offered patients a means for theft and violence, see below excerpt:

Participant 2:
“Patient Safety” meaning? Like will
they take a swipe at me while I’m using
them or that they would steal it from
me?

Losing a computer could adversely affect workflow in a busy clinic dependent on technology. The notion of loss-by-theft was mentioned by another participant:

Participant 1:
At one point, we were thinking of
having these [devices, but we
considered:] theft, dropping...and
having considered this, but then we
thought well, we’d go with this

[desktop computer], because this is more secure.

The app ordinarily used the computer network to access patient data. This was simulated in the training mode of the app, by using artificial data-retrieval delays during the session. One participant commented that using cloud-based patient records could negatively affect patient confidentiality. This, by having a potential for unauthorized access by outside parties who acted with malicious intent:

Participant 1:
Personally, I don't like the cloud or any of these things because it's exposing you, the patient, to easy hacking.

Not all participants focused on antisocial behaviour when considering device-use and safety. One participant considered the dynamic attributes of data displayed on an electronic screen, compared with the static nature of printed paper. The ability to colour-code the data to indicate important facts, was perceived as an improvement to patient safety:

Participant 6:
Good point so as long as you have normal values in colour or in some way indicated and highlighted? Any device really, so if you have that [feature] like our EMR does then the chances that something would be missed is reduced.

Also seen as a step towards improving patient safety, was the concept of utilizing intelligence, as supplied by the computer, to improve decision-making. More fundamentally, the ability to validate data on entry which could not be automatically accomplished in a paper-based system; this was perceived as safer than paper as described by participant 1:

Participant 1:

Thing about this though is that you've got the ability to build in [validation]. So you could say "so this is a value outside the limits, did you mean it" or "the last two patients we gave this to died" or something like that. So your ability to build in fail safes is higher with the devices than it is on paper.

One participant commented that the layout that collated data from what would normally be sourced from many places, i.e., several different sheets of paper, improved safety because the relationships between results can be seen and interpreted if values were presented in the same view:

Participant 9:

Because they gather all the info in one place, quickly ... [the app adds] to the safetybut because they are technology, they have some difficulties working with other technologies, you need to be familiar with the options and the functions, and you need to be patient and not get frustrated and put it away. But generally, I think it adds to patient safety.

Another participant cautioned that the iPad® did not show enough data on a screen to ensure patient safety and reinforced the serious consequences of mistakes made with the patients' data as described by a participant below:

Participant 1:

The amount of information you're getting on the iPad® with this app is minor. I wouldn't make a diagnosis based on that. It's just not enough. I'd need more [information] before prescribing drugs. I could kill you just as soon as I could cure you.

Likert Results: iPhone® and iPad® Compared

This study utilized Likert questionnaires, all of which used a nine-point scale to indicate participants' agreement or disagreement with each question: where 1=strongly disagree, 5=neutral, and 9=strongly agree. Most questions were stated in the context of the iPhone® or the iPad®, except for two questions that were stated in the context of the app itself.

The participant session had two periods of device-use each, followed by a brief interview and a questionnaire about the device used. The session ended with a single questionnaire which contained statements about both devices. One final question asked the participant to choose a preferred device. The responses were collected from each participant and categorized into iPhone®, iPad®, and app. These sections contained visualizations and analyses of the data.

Likert Scores. Each chart in Figure 21 and Figure 22 contains the results of two similar statements from each device. As an example, Figure 21 shows a pair of charts; in the first chart, the two statements were: *"I would avoid using the iPhone® in clinical practice,"* and *"I would avoid using the iPad® in clinical practice."* For the chart's title, the two questions were merged and condensed into *"I would avoid using the device in clinical practice."*, where the device is shown as iPhone® (open circles) or as iPad®

(solid circles). Refer to the appendix for the exact wording of each question. The remainder of the charts that follow in Figure 22 are handled in a similar manner.

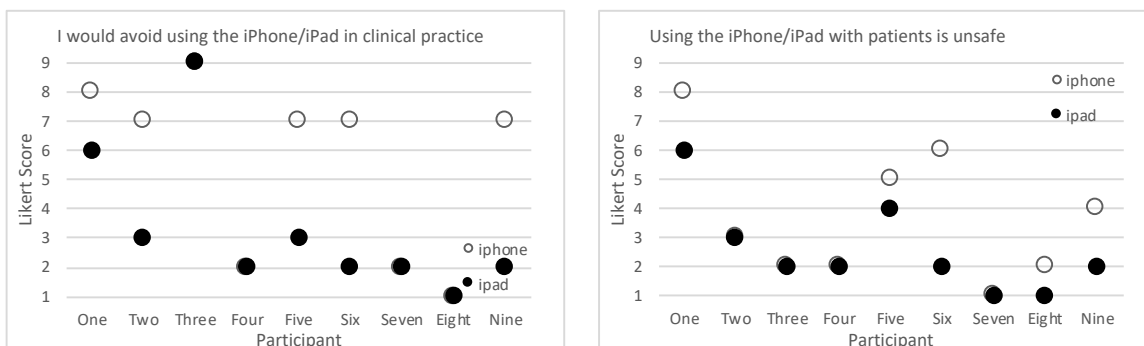


Figure 21: Likert result charts device perceived safety

Two Likert results concerning participants' negative perceptions of using the mobile device in clinical practice. The charts indicate participants' Likert responses to the statement indicated as an outlined circle for the iPhone® and a solid circle for the iPad®.

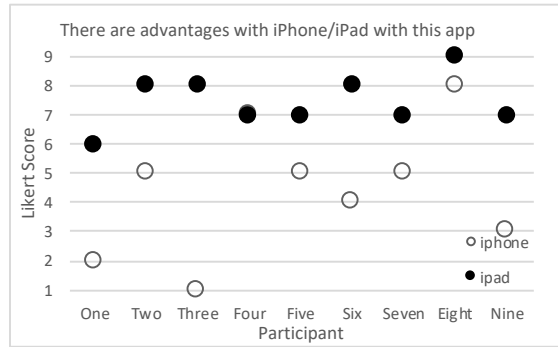
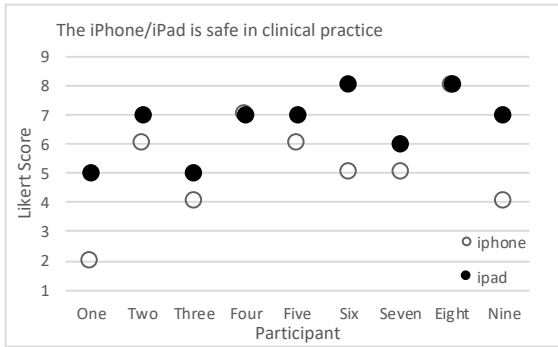
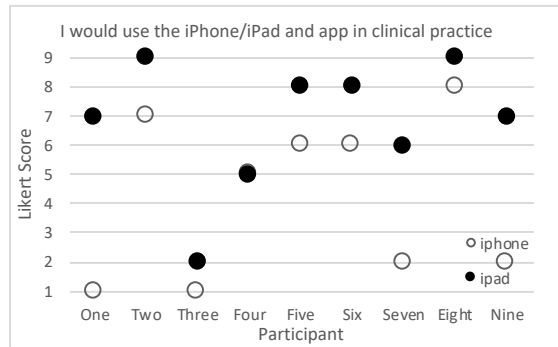
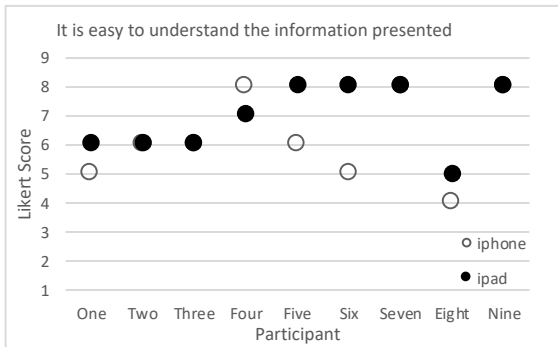
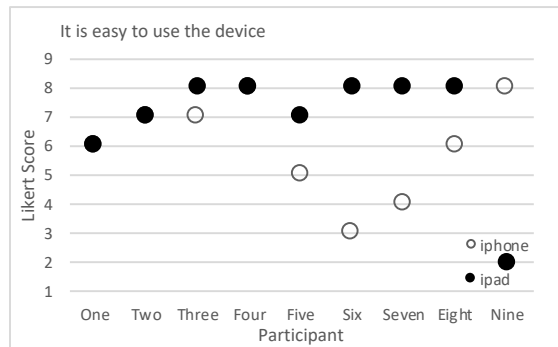
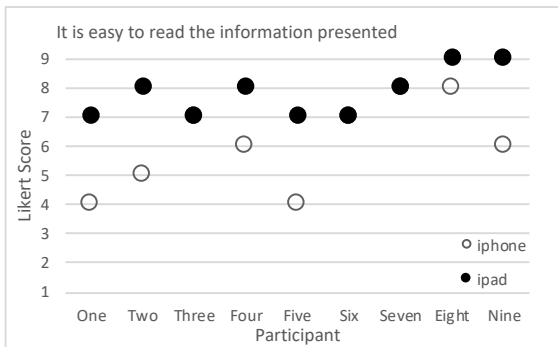
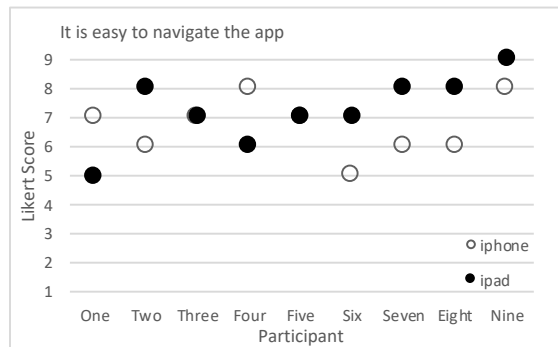
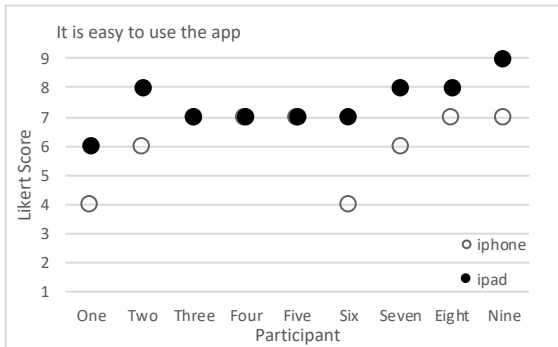
The charts in Figure 21 show a separation between responses concerning the iPad® and the iPhone®. The questions had required the participant to consider the device in a negative context. Participant Three's response to the statement "I would avoid using the device in clinical practice" was an outlier. This participant had decided a priori that neither device was to be used; he or she had indicated that no computer was used during patient encounters, and therefore the response indicated that neither the iPhone® nor iPad® was to be used in clinical practice. Participant Five responded to the second statement strongly indicating that the devices were not perceived to be unsafe. One could infer from these responses, without certainty, that if Participant Three's personal policy had allowed computer-use in the presence of a patient that the response to "I would avoid using the device in clinical practice," may have been different.

The majority of participants, i.e., six of nine, indicated agreement that they would avoid the iPhone® in clinical use, and that they disagreed with the same statement about

the iPad[®], i.e., participants would not avoid using the iPad[®] in the clinic. Participant One felt that both devices were to be avoided in the clinic. The iPad[®] is less so than the iPhone[®]. Seven participants indicated the iPad[®] should not be avoided and three participants indicated the iPhone[®] should be avoided. This may infer the participants perceived the iPhone[®] to be either less safe or less useful than the iPad[®].

Responses to the statement “Using the device with patients is unsafe,” shown in Figure 21, indicate that the majority of participants did not perceive either the iPad[®] or iPhone[®] to be unsafe when used with patients. Two participants indicated that the iPhone[®] was unsafe; and, one, that the iPad[®] was unsafe. The majority perceived both devices to be safe, the iPad[®] more so than the iPhone[®].

The charts in Figure 22 visualize responses from all participants on statements that demonstrate a preference for either the iPad[®] or the iPhone[®]. The trend shows that the iPad[®] was preferred. The solid-circled iPad[®] markers in the charts were generally above the open-circled iPhone[®] markers more often than they were below.



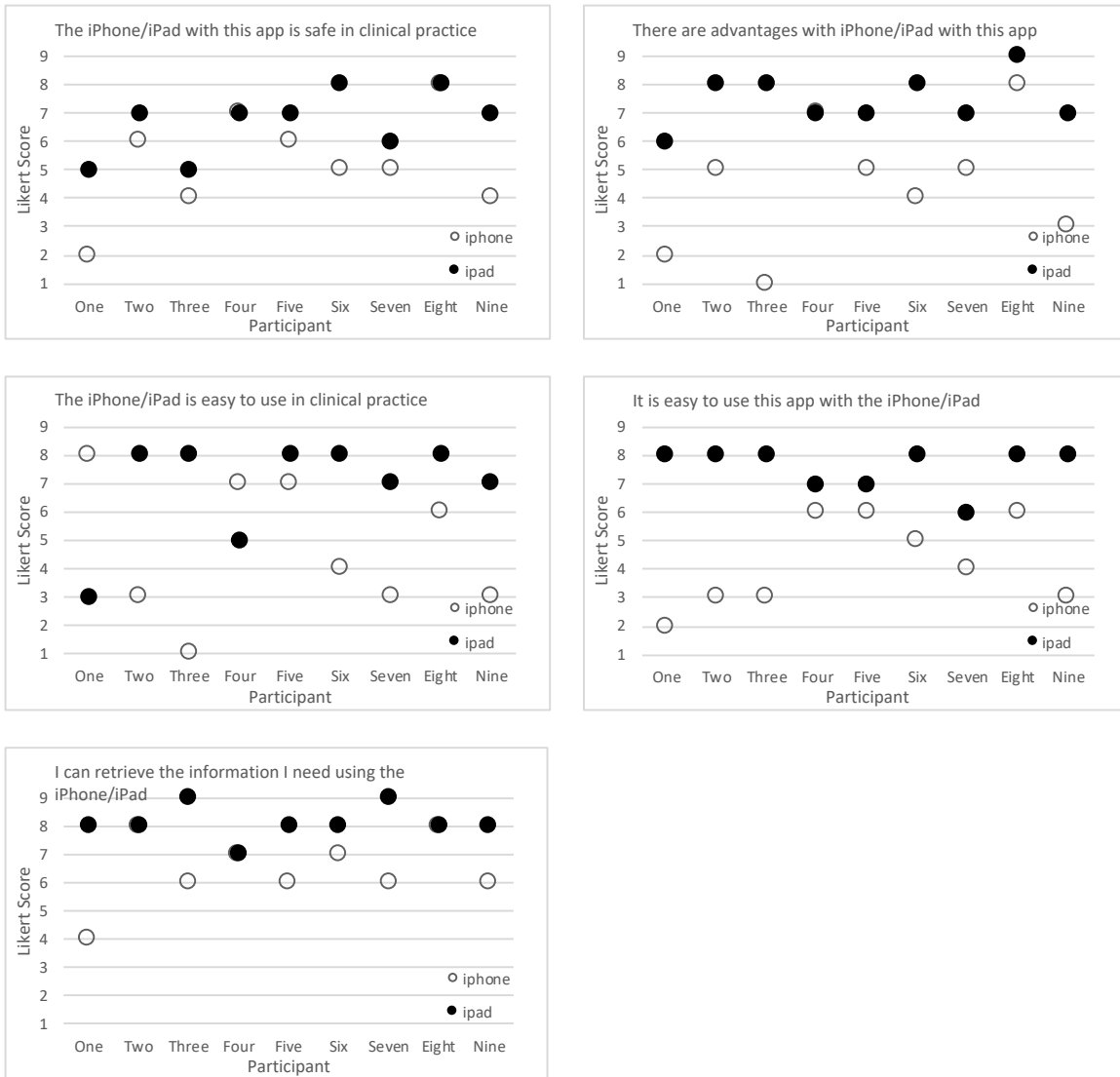


Figure 22: Likert responses for remaining statements

The remaining statements comparing participants' perceptions of iPad® and the iPhone®. The charts indicate participants' Likert responses to the statement indicated as an outlined circle for the iPhone® and a solid circle for the iPad®.

Participants were also challenged with two Likert statements, shown in Figure 23, which referred solely to the app without reference to a device: “This app is not safe for use in a clinical setting,” and “This app is safe for use in clinical team meetings.” The participants’ responses to these questions indicate the app is safe for “clinical use” but not safe for “clinical team meetings.” Participant One’s response to the first question seems to be an outlier. The cohort is in general agreement with all other responses, but in disagreement with the statement. In the second statement, participant One’s response was consistent with the group’s response. Given that the first statement is expressed as a negative, it’s likely participant One’s response was made against a positive statement and should be a two rather than an eight.

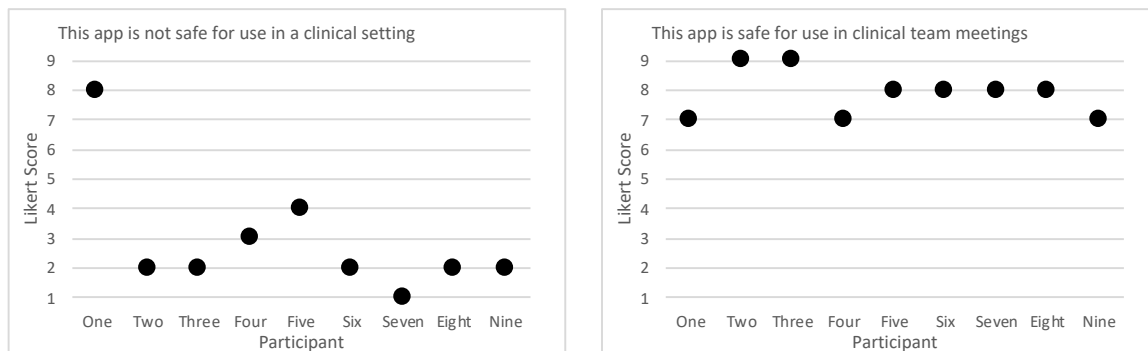


Figure 23: Likert responses to statements concerning the app

These statements do not reference either device. They are intended to focus participants on the app itself.

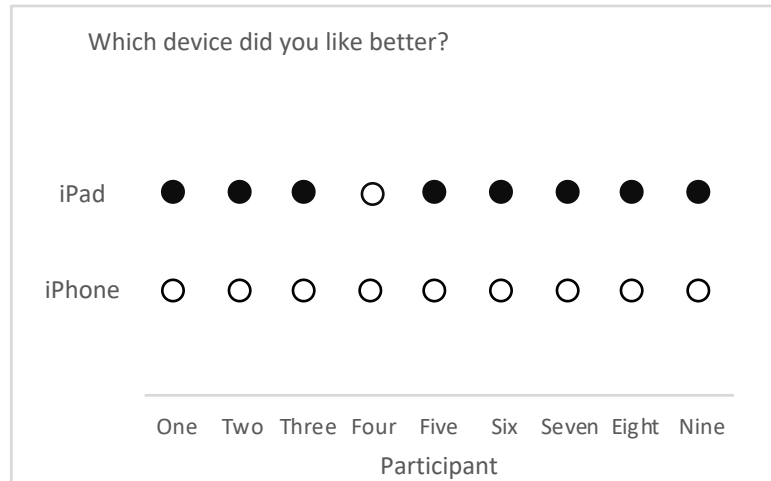


Figure 24: Results for overall device preference

One final question was presented to participants at the close of the session. The result is shown in Figure 24. The final question resulted in a preference for the iPad[®] by 100% of participants who responded. Participant four refused to answer this question, leaving the following comment “Depends on the use!” This participant, who was argumentative, seemed to favour the iPad[®] during the sessions and the sum of all his or her Likert responses were: iPad[®]: 81, iPhone[®]: 84, suggesting a neutral view consistent with the response given to the final question.

Summarized Likert Scores

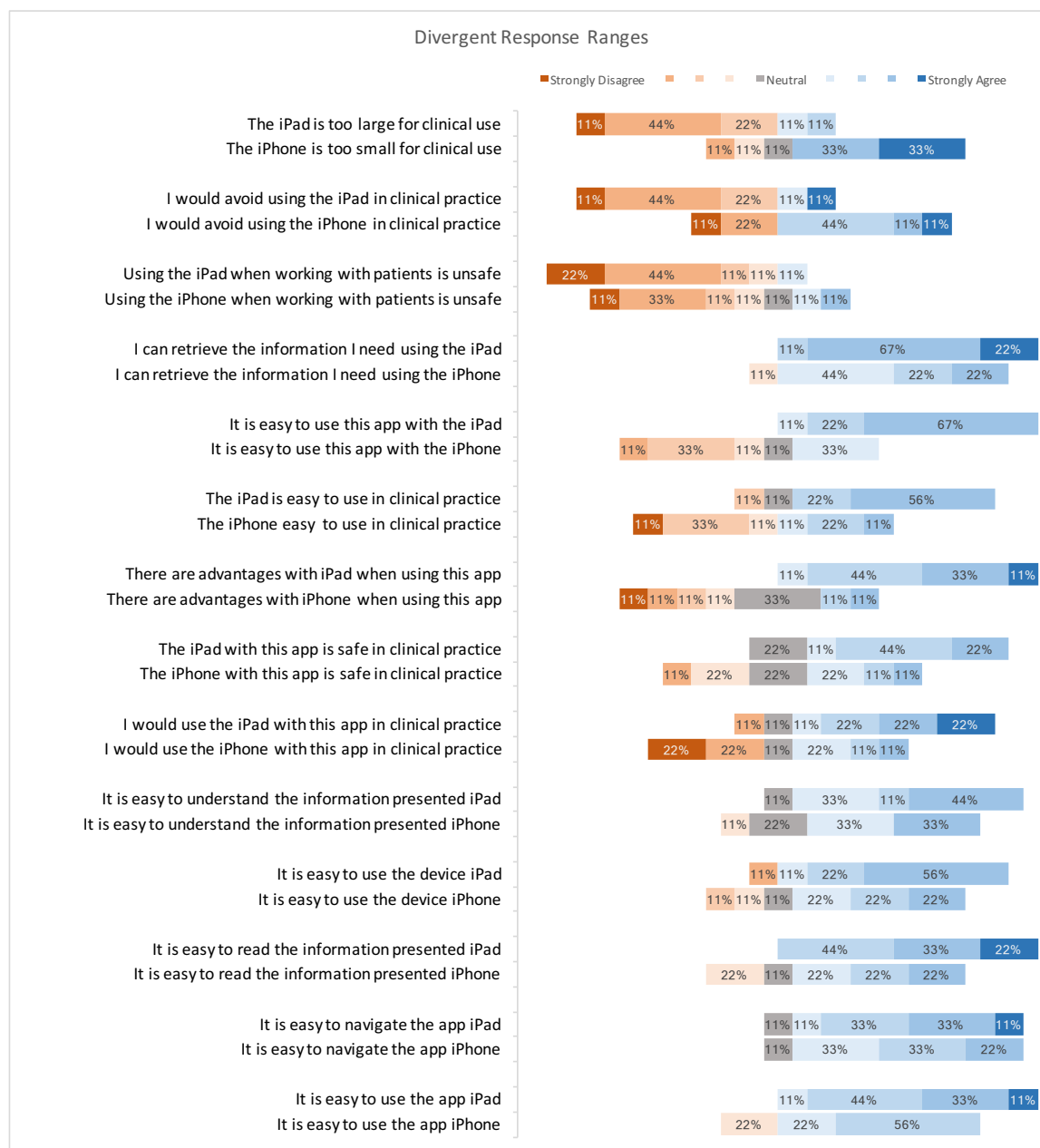


Figure 25: Likert responses for all participants' responses

This chart shows an overview of all Likert responses as a distribution for each statement in the questionnaires. Statements were presented twice, once for the iPad® and once for the iPhone®. Responses are shown in pairs of segmented bars. Using the topmost pair as an example: “The [iPad®/ iPhone®] is too [large/small]...”: the responses for the iPad® are shown in the first bar; in which 11% of participants strongly disagreed (far left), none were neutral and 11% were halfway between neutral and Strongly Agree. The second bar shows the iPhone® results. Comparing the two bars in this pair indicated a disagreement for the iPad® statement, and an agreement for the iPhone® statement, and, that neither statement had unanimity.

Figure 25 shows the range of responses to all iPhone®/iPad® statements from all participants. The statements are grouped to have the iPad® statement above the iPhone® statement in each pair. Each statement is followed horizontally by a bar with many segments. Each segment in the bar shows the percentage of participants that responded to the statement with a particular Likert score. For example the first question, “The iPad® is too large for clinical use,” 11% of participants responded with “Strongly Disagree;” the second statement “the iPhone® is too small for clinical use,” 33% of participants responded “Strongly Agree”. Colour codes for each segment represent the steps which make up the Likert scale.

The results are the same as those of the Figure 22, a consistent preference for the iPad®. The positive-worded statements exhibit a shift to the right for iPad® responses, and a shift to the left for the negative-worded statements. Both these shifts indicate a preference for the iPad®. One notable exception is the statement, “It is easy to navigate the iPhone®/iPad®,” is neutral in that there is no clear preference and only the distribution of responses differs. It is easy to see that the median and mean responses are also identical.

The questions shown at the top of Figure 25: “The iPhone® is too small for clinical use,” and “The iPad® is too large for clinical use,” could not be compared directly in the raw results, as the wording differs within the pair; neither can they be compared here to indicate a preference for one or the other device. The result of this pair is that participants thought the iPhone® was indeed too small for clinical use, and that the iPad® was not too large for clinical use.

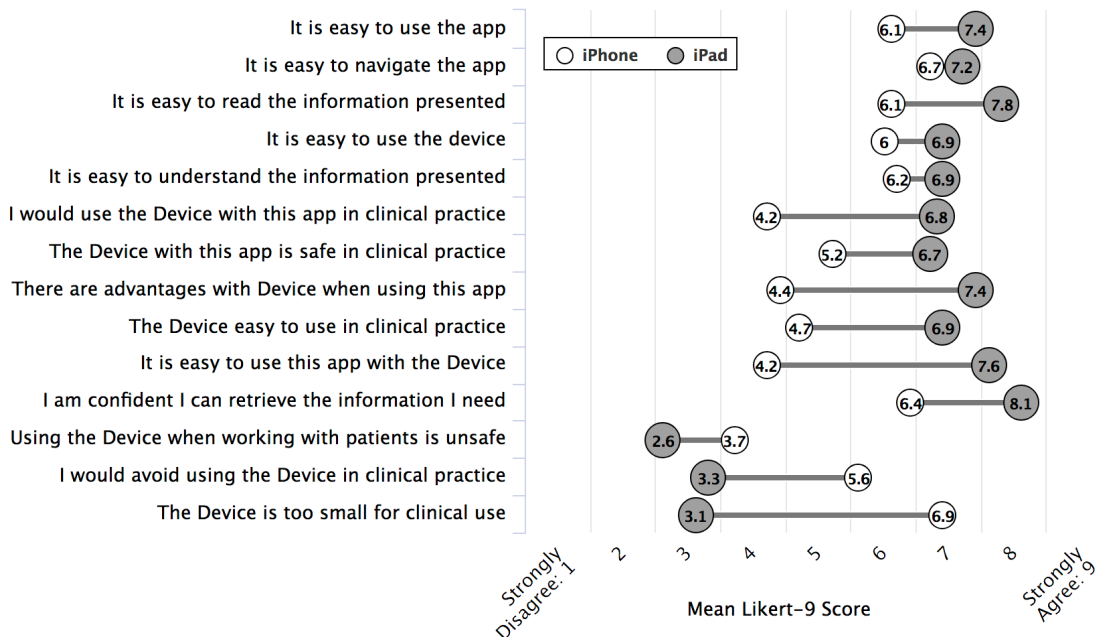


Figure 26: Mean Likert Scores for each question

This chart shows the range of mean response values for Likert statements between the iPhone® (open circles) and iPad® (closed circles). Where the two circles are far apart there is disagreement between the responses for different devices. For example, the mean response to: “It is easy to understand the information presented,” was similar for the iPad® and the iPhone®, i.e., mild agreement. Conversely the mean response to “I would use the device with this app in clinical practice” was in disagreement between the iPhone® (disagreement) and iPad® (agreement).

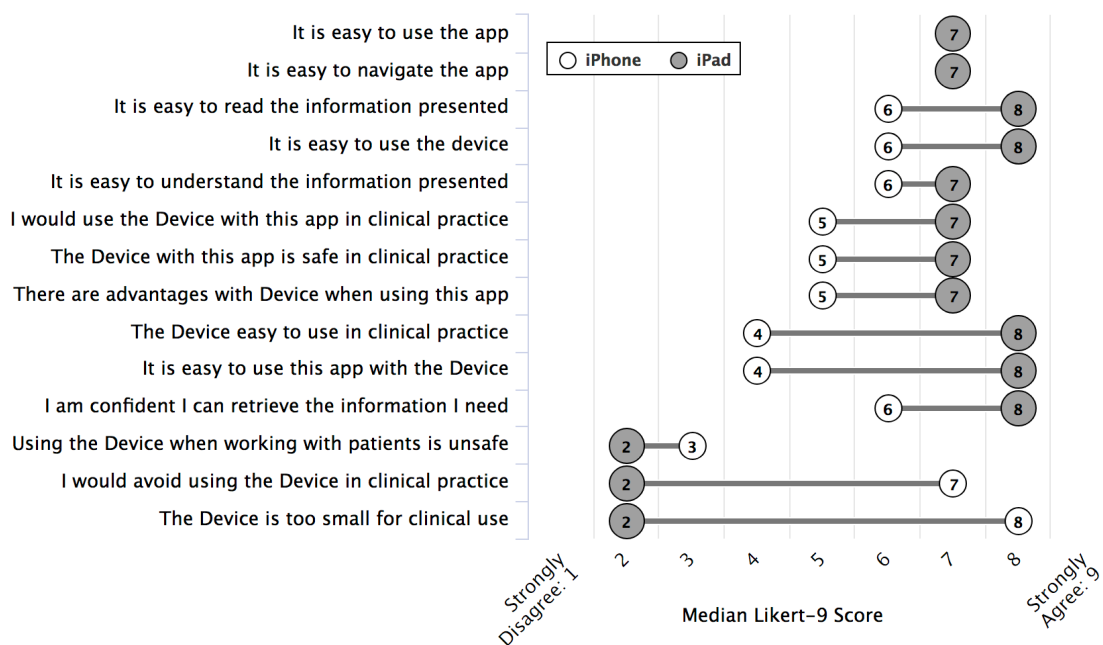


Figure 27: Median Likert Response for each question

This chart shows the range of median response values for Likert statements between the iPhone® (open circles) and iPad® (closed circles). Where the two circles were far apart there was disagreement between the responses for different devices. For example, the median response to: “It is easy to use the app,” was identical for both the iPad® and the iPhone®, i.e., moderate agreement. Conversely, the median response to “The device is too small for clinical use” had disagreement between the iPhone® (it was too small) and iPad® (it was not too small).

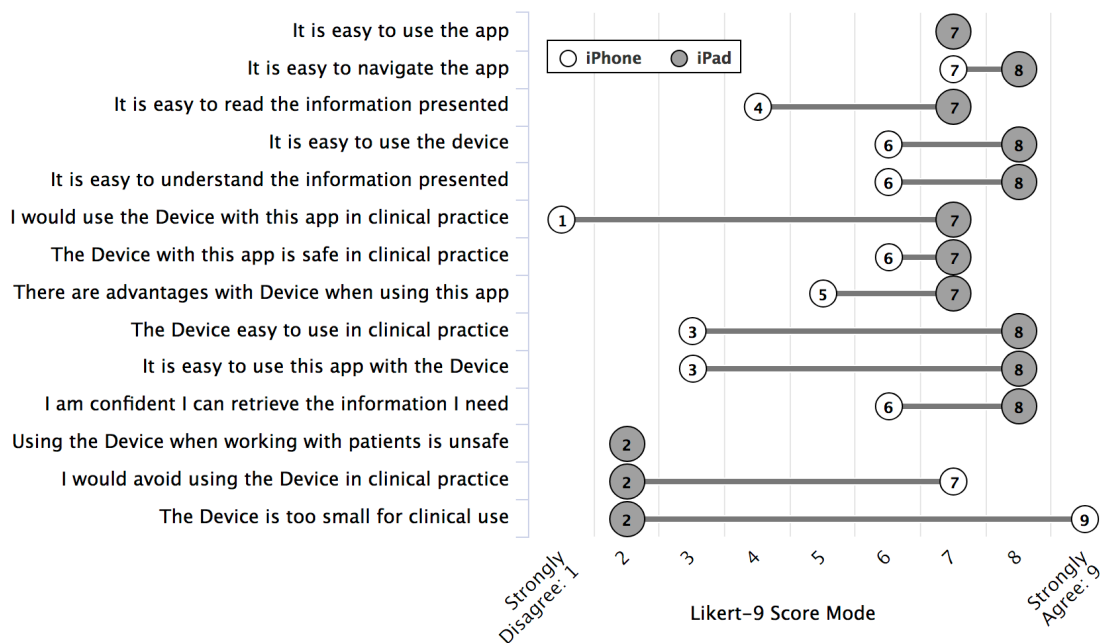


Figure 28: Mode of Likert responses for each question

This chart shows the range of mode of response values for Likert statements between the iPhone® (open circles) and iPad® (closed circles). Where the two circles are far apart there is disagreement between the responses for different devices. For example, the mode of responses to: “It is easy to use the app,” was identical for the iPad® and the iPhone®, i.e., moderate agreement. Conversely the mode of responses to “The device is too small for clinical use” was in disagreement between the iPhone® (the most frequent response was moderate disagreement) and iPad® (the most frequent response was strong agreement).

The results charted in Figure 26, Figure 27, and, Figure 28 shows the trend of stronger agreement for iPad®-related questions and stronger disagreement for iPhone®-related questions. The question “*The iPad®-too-large / iPhone®-too-small for clinical use*” has the widest range across the three reported results. This question’s response indicates a strong agreement with the concept that the iPhone® is not usable for clinical use. The ease-of-use and ease-of-navigation questions are neutral with an agreement for both devices.

Table 7 summarizes the responses for each Likert statement for each device. Each question’s response indicates a preference for the iPad® over the iPhone®. The scores tabulated are normalized to a scale of one to one hundred where one indicated the strongest disagreement, 50 indicated neutrality and 100 the strongest agreement. The difference between the two scores is used to determine a relative preference between the two devices. Each point on the Likert scale normalizes to a value of 11. Thus, a difference of between 22 and 44 indicated a moderate preference, and, a difference greater than 44 would have indicated a strong preference, though none were found. Some results of note from Table 7 were as follows:

For the statement “*It was easy to navigate the app,*” all participants agreed on both devices with the iPad® being somewhat preferred. The statement “*It is easy to understand the information presented,*” had the least difference between devices. The score was 77 for the iPad®, and 69 for the iPhone® both indicating moderate agreement. The statement with the highest score for both devices was “*I can retrieve the information I need using the device,*” with 90 – indicating strong agreement – on the iPad® and 72 – indicating moderate agreement – on the iPhone®. The statement with the lowest score for

both devices was “*I would avoid using the device in clinical practice*” with 62 for the iPhone® and 37 for the iPad® indicating agreement the iPhone® would be avoided and the iPad® would not be avoided. Responding to “*Using the device with patients is unsafe*” the scores indicate that the iPad® was preferred but that both devices were perceived as being safe to use with patients.

All questions indicate a preference for the iPad®. The total of all the scores indicated a preference for the iPad® by a margin of 178 points, or, a ratio of 1.23:1.

Table 7: Summary of Likert responses including total scores

Statement	Percentage Likert Score [⊠]	Δ	Participants Responses indicated that:
It is easy to use the app	<i>iPad</i> [®] 83	15	All participants agreed. The iPad[®] somewhat preferred
	<i>iPhone</i> [®] 68		All participants agreed.
It is easy to navigate the app	<i>iPad</i> [®] 80	6	All agreed. Neither device significantly preferred
	<i>iPhone</i> [®] 74		All agreed.
It is easy to read the information presented	<i>iPad</i> [®] 86	18	All agreed. The iPad[®] somewhat preferred
	<i>iPhone</i> [®] 68		6 agreed, 1 neutral, and, 2 disagreed
It is easy to use the device	<i>iPad</i> [®] 77	10	All agreed. Neither device significantly preferred
	<i>iPhone</i> [®] 67		5 agreed, 1 neutral, and, 2 disagreed
It is easy to understand the information presented	<i>iPad</i> [®] 77	8	8 agreed, 1 neutral. Neither device significantly preferred
	<i>iPhone</i> [®] 69		6 agreed, 2 neutral, and, 1 disagreed
I would use the device with this app in clinical practice	<i>iPad</i> [®] 75	28	7 agreed, 1 disagreed, and, 1 neutral. iPad[®] preferred
	<i>iPhone</i> [®] 47		4 agreed, 1 neutral, and, 4 disagreed
The device with this app is safe in clinical practice	<i>iPad</i> [®] 74	16	7 agreed, 2 were neutral. iPad[®] somewhat preferred
	<i>iPhone</i> [®] 58		4 agreed, 2 were neutral and 3 disagreed.
There are advantages with the iPad [®] /iPhone [®] and this app	<i>iPad</i> [®] 83	34	All agree, the iPad[®] preferred
	<i>iPhone</i> [®] 49		1 agreed, 3 were neutral, and, 4 disagreed
The device is easy to use in clinical practice	<i>iPad</i> [®] 77	25	7 agreed, 1 neutral, and, 1 disagreed. iPad[®] preferred
	<i>iPhone</i> [®] 52		4 agreed and 5 disagreed
It is easy to use this app with the iPhone [®]	<i>iPad</i> [®] 84	37	All agreed. iPad[®] preferred
	<i>iPhone</i> [®] 47		3 agreed, 1 neutral, 5 disagreed
I can retrieve the information I need using the device	<i>iPad</i> [®] 90	18	All agreed. The iPad[®] somewhat preferred
	<i>iPhone</i> [®] 72		8 agreed, 1 disagreed.
I would avoid using the device in clinical practice*	<i>iPad</i> [®] 63	25	2 agreed and 7 disagreed. The iPad[®] not avoided
	<i>iPhone</i> [®] 38		6 agreed and 3 disagreed. The iPhone[®] avoided
Using the device with patients is unsafe*	<i>iPad</i> [®] 72	13	1 agreed, 1 disagreed, and, 7 disagreed.
	<i>iPhone</i> [®] 59		2 agreed, 1 neutral, 6 disagreed. iPhone[®] somewhat less safe.
Sum of Likert values:	<i>iPad</i> [®] 951		
	<i>iPhone</i> [®] 773		
	Δ 178		

* response score was inverted for consistency because the question was phrased negatively

⊠ Scores are the sum of responses normalized to a value in a range of 1-100

Examining the aggregate of all statement responses together the mode, i.e., the most common Likert value, in the iPad® context was an eight, indicating a moderate agreement. The mode for the iPhone® was a six indicating an agreement of less strength and closer to neutral than that of the iPad®’s responses. Median and mean responses also result in a similar preference for the iPad® and are detailed in Table 8. Both the median and mean score for iPad®-related statements were greater than the median and mean scores for iPhone®-related statements

Table 8: Aggregate median, mean, and, mode response values of all Likert statements

Device	Median (MAD)	Mean (σ)	Mode
iPad®	7 (1.0)	6.6 (2.12)	8
iPhone®	6 (1.0)	5.4 (2.08)	6

Table 9: Preferences for all Likert Statements

Device	n (of 117)	Percent
Preferred iPhone®	16	14%
Preferred iPad®	71	61%
Neutral	30	26%

For all individual responses, 61% indicated a preference for the iPad®, compared with 14% for the iPhone®, and, 26% neutral, see Table 9. The count of responses which preferred the iPad® were more than 4 times (4.43x) the count of responses which preferred the iPhone®. In addition, responses preferring the iPad® were more than 1½ times (1.54x) the neutral and iPhone®-preferred responses combined.

Chapter 6: Discussion

The purpose of this work was to explore the effect of screen size of two mobile devices on the usability, usefulness and safety of a healthcare software app common to both devices. What resulted was not only the identification of some key items for aspects of usability, usefulness and safety, but also more general themes about integration into clinical workflows. Using the literature reviews to supplement the data collected in the study produced on this topic a wealth of information in several areas. This chapter combines the research findings and the review findings in order to discuss and answer the three original research questions. General findings from this work are discussed below, followed by a discussion addressing each of the original questions posed at the beginning of this thesis.

General Findings

The following section highlights a few of the key results of this research in the areas of usefulness and safety of handheld mobile devices in clinical work with patients.

Skill and knowledge. Participants were aware that the app required skill to use efficiently but were not concerned that learning this skill would be difficult. This perception was reinforced by the lack of negative comments about the app's learning curve. The dexterity needed to interact with the user interface was a factor, as was the need to physically navigate to discover patient data. Most demanding of the user's interaction skill was the reports section on the iPhone®. This section featured a built-in PDF viewer that allowed participants to scroll, pan, and, zoom – often all at the same time – to bring the virtual printed page into legibility. The physical interaction was

perceived to have imposed a time penalty. The participants thought charting using the iPhone® would take longer than doing so on the iPad®. The participants who commented about this time penalty did not make explicit comparisons on the time needed to shuffle through paper records or browse an EMR to perform the same task.

One participant, who worked with geriatric patients, expressed a negative view of using the iPad® with older patients. The concern related to the potential confusion some older patients may experience with the new device. This hesitation to adopt new technology in encounters with older patients, would be mitigated over time as the technology matured – when patients familiar with the mobile devices will become aged patients. In the meantime, the participant advocated implicitly for a more technologically conservative approach, i.e., paper, in encounters with older patients. There are two points of discussion:

First: The notion, expressed by some participants that older adults, both patient and physicians, are more likely to become hampered by the increasing computerization of daily activities. During this research, there was anecdotal evidence to refute this argument. Some participants in this study were older and some were younger. The eldest participant had had the longest experience with mobile handheld devices and handled them with the greatest apparent ease. The youngest participant, who was less experienced, handled the devices with less ease and with greater profanity. To judge ability and acceptance so broadly on the basis of age provides a disservice because:

Second: Mobile handheld devices are adaptable. They can be configured to enlarge the text size for users with poor vision making content more readable than on paper. Due to their portability, the iPad® included, they may be used by clinicians at the bedside as a

visual aid in a more personable manner than a desktop computer. There are a great many potential advantages to utilizing technology in this area; however, further discussion is beyond the scope of this thesis.

The notion that experience is the arbiter of ease-of-use is reinforced by a study by (Miller et al., 2017). The only predictor they found for a patient needing assistance with a device was a lack of experience with the device. The perceived difficulty in adapting to technology is related to experience with the technology and not the age of the user. Judicious implementation of mobile device technology could improve the delivery of care to those patients who could perhaps benefit more than most from its malleability.

Despite the problematic user-interface and the negative comments about some aspects of the iPad®'s use, the foregoing participant felt that these problems are not significant when contemplating the wealth of information mobile devices contain or to which they have access, on a device that fits into one's pocket.

One participant provided a philosophical perspective, implying that the iPad® is a boon for knowledge-based work:

Participant 9:
[The iPad® and iPhone®] weren't
difficult ... I come from an age in
which had to memorize most of this
information, so having something this
size ... is a miracle ... and it's
worthwhile.

Another participant recalled that during training as a physician, they were required to commit a great many facts to memory. On using the reference ranges shown in the lab results section of the app, this participant had the insight that some of this memorization is redundant now; because a device can provide the information at a tap of a button. Could this lead to clinicians' dependency upon mobile devices for accurate healthcare?

Trust that information supplied by technology is infallible, is misplaced trust. It is said that computers do not make mistakes; rather, people make mistakes. Software and the devices it runs on are increasingly complex. Blindly assuming that the system will behave the way that the user expects to is an error of judgement.

The doctor-patient experience. One participant expressed concern that the use of either device during an encounter with a patient could have a negative impact on their workflow. Another participant was concerned that the device may impede communication. None of the participants stated that they used their own devices in medical practice, so this concern is not likely due to their personal experience with the devices. All but one participant used a computer with patients leading the researcher to speculate that the perceived negative impact on their workflow arose not from their personal experience with the mobile devices, but rather from their experience with computers in the clinic. The notion that mobile devices would impede interpersonal communication goes against evidence found in the literature, that shows mobile devices afford new ways to permit body language to improve doctor-patient encounters (Alsos et al., 2012). One participant commented that the iPad® would be more conducive to a bedside visit than a ‘wired-in’ computer. The participant who thought the iPad® would be an encumbrance in a patient encounter did not necessarily lack insight. More likely they were highly concentrated on using an unfamiliar device to complete the allotted task. Given practice to gain the skills necessary to use the app efficiently, they would be more at ease and more able to concentrate on the patient during an encounter.

Participant 4:

Given that people use laptops or even monitors ... I mean, will the iPad® get in the way of you talking to your

patient? Well, no more than the paper you were writing on in the first place. It might be better if you are having to explain things to a patient, the iPad® might be a better thing to do it with.

This is a notion grants equivalence to whatever technology the physician uses to contain patient-related information, i.e., a clipboard with a chart, a set of paper notes, an iPad® – all being an object used for the patient by the physician. One participant made a comment on the iPad® that suggested the device could have an impact on their workflow, by allowing data entry and charting to be performed locally to the patient. A case study by Nolan, (2008), tells the story of a medication error in which a physician using a remote computer ordered medication for the wrong patient. The medication order, a paralytic drug, was dispensed and administered with life-threatening results when the patient experienced the effects of the paralytic while walking away from the hospital room.

This situation may not have happened if the order had been created with a mobile device at the bedside because the context of the order would have been preserved. Nevertheless, the danger of an error in prescribing still exists because the mobile device may be taken into a room with the clinician and the order could be created away from the patient just as easily as it was for the patient in the case study.

Display Size and Perceived Safety

The first question posed at the start of this work was: “Is there a correlation between the size of the handheld device’s display and the perceived safety of the iPhone® and the iPad®?” One participant thought that the existing technology, a computer in the examination room, causes problems with safety and confidentiality during patient

encounters because the computer display is visible to the patient. The same could be said about a mislaid paper chart, the problem manifested, if the previous patient's paper chart or computer screen is displayed when the next patient arrives, sits, and has a clear view of the private information. Unlike the paper chart, the electronic displays (ie. iPad®, or computer monitor), can be set to hide their screens after a brief period. As a result, electronic displays have an advantage over paper charts from this point of view.

Some participants perceived the benefits to patient safety when technology is employed to improve it. Participants considered the dynamic attributes of data displayed on an electronic screen compared with the static nature of printed paper. The following three innovations were perceived by participants as things that may improve patient safety:

First: employ the mobile devices' display to colour-code the data to indicate important facts. This was demonstrated in the app where out-of-range values are highlighted in a colour meant to draw the user's attention.

Second: utilize intelligence, as supplied by the computer, the system could be programmed to improve decision-making, 'Intelligence-based decision support', for example, as one participant expressed it:

Participant 1:

So, you could say "so this is a value outside the limits, did you mean it", or, "the last two patients we gave this to died", or something like that. So, your ability to build in fail safes is higher with the devices than it is on paper.

Third: Permit the user to customize layouts by rearranging sections that allow a physician to collate data most relevant to a specific clinical workflow. From a healthcare usability perspective; however, this feature idea could have the effect of negating any

usability metrics by allowing the user to make changes to the UI that may not have been accounted for when initially tested.

Diagnostic imagery was available in the app on both devices however, the app was not designed for diagnostic imaging purposes. Two participants commented, with the conclusion, that they would not use either device with imaging other than as a visual aid when in consultation with patients. That stated, there was an inevitable urge to quickly consult the mobile device due to its portability, as had happened during the study to look up ancillary information about the spider chart shown. It is plausible that viewing imagery on mobile devices can inform or affect the practitioner's decision-making process.

No participant agreed that this particular app was safe in a clinical setting. However; they did agree it was safe to use during team meetings. The disparity is explained by results from a paper found in the literature that concludes 'different people use the same tool in different ways.' The data on the device would be the same in team meetings as it would be in a patient encounter; therefore, the trustworthiness of the app is perceived to be acceptable. If participants did not like the app, they would not use it at all, but, it would be used in team meetings. Therefore, the app is perceived to be useful. The consequences of using the app would be the same in both situations, if meetings were held to determine how to provide care. If so, then the app would probably be avoided with a patient because it is awkward to use and would adversely affect the doctor-patient experience. On the other hand, the more likely scenario is that team meetings would be more general in scope and focussed on all the teams' patients rather than a single one. With this in mind, the app was perceived to be useful to participants as a source of more

general information, and perceived to not be useful as a tool for more detailed analysis of a single patient. The trust eroded by the untrustworthy demo data may have contributed to this perception.

One subtle safety issue was discovered by a sole participant on the iPad®. This participant discovered that additional information about the patient's allergies was available in the Allergies section when the edit mode was toggled on. When in the toggled-on state, more information about the patient's allergy was displayed. Clearly a user should not need to switch modes to read more data.

Other participants used design flaws to work around usability; for example, the use of the 'close' button in place of a 'back' button. Once users discover that important data are available in edit view, they would most certainly make use of this flaw and use it as a workaround. Requiring the user to enter an editable state, the UI increases the likelihood the user would inadvertently change the data. In this case, the app user in this study was a client for an EMR. Editing the data in the app would change the data in the EMR. The next clinician to use the patient's record would see the altered data and would likely trust it even though it had been changed in error. Allergies can cause life-threatening conditions; inaccurate allergy data could have life threatening consequences as described by one participant:

Participant 1:
I'd need more [information] before
prescribing drugs. I could kill you
just as soon as I could cure you.

Specifically, the edit-to-view issue discovered in Allergies is: a defect in the software, a usability problem, a safety issue, and a good example of the root cause of a technology-induced error. Generally, the design decision to make the app read-only, except for

explicit cases of editable data, improved the safety of app. Separating the GUI views into editable and non-editable modalities is a benefit to safety, however, important clinical data that exist only in the editable view is potentially dangerous.

Participants also expressed mixed perspectives regarding patient safety. They did not perceive the device as an impediment to communication. They also thought the device, through which they could have rapid access to patient data, would improve safety by having data at hand. One participant suggested that the app could be improved to provide decision support by automatically searching for information related to their patient's condition and treatment. Given that potential, participants warned that the current app had not gone far enough with decision support and the devices were not large enough to display enough information with which to diagnose the cause of a patient's problems. The iPhone[®], with its smaller controls and fonts, was seen as too small and difficult to manage for this task.

Despite no participants agreeing that the app was safe to use in a clinical setting, responses to the statement "Using the device with patients is unsafe," shown in Figure 21, indicate that the majority of participants did not perceive either the iPad[®] or iPhone[®] to be unsafe when used with patients. The majority of participants perceived both devices to be safe, the iPad[®] more so than the iPhone[®], but not the app itself. Yet, the majority of participants, i.e., six of nine, would avoid the iPhone[®] for clinical use, and they would not avoid the iPad[®] in the clinic. On the subject of screen size, participants agreed that the iPhone[®] is too small for clinical use. On the final question: "Which device did you prefer?" no participant chose the iPhone[®]. The results indicate the iPhone[®] was not significantly more difficult to use or navigate. Therefore, the iPhone[®] would be avoided

in clinical practice by the participants because of its size, not because they thought it was unsafe or too difficult to use, and despite that level of comfort, the app would be avoided.

Responses for all safety related Likert questions, such as, "the device with this app is safe in clinical practice." There is a preference for the larger screen. Two of the questions were expressed as negatives, "I would avoid using the device in clinical practice" and "Using the device when working with patients is unsafe." Participant responses to both these Likert Scale items also indicated a preference for the larger device. Two of the questions were expressed as negatives, "I would avoid using the device in clinical practice" and "Using the device when working with patients is unsafe." Both questions indicate a preference for the larger device. The device with the larger screen is perceived to be safer.

In the responses to all other Likert questions, there is a preference for the larger screen. Quantitative results indicated that the larger screen had 8% fewer usability problems than the smaller screen. Lastly, the overall preference, 'which device do you prefer,' as indicated by all respondents, was the iPad®.

Users committed 55% more Safety errors on the smaller screen compared with the larger screen. The participants indicated a strong preference for the iPad® and the results of coding their behaviour indicates that the iPad® was the safer device.

The iPad® is perceived to be safer relative to the iPhone® when used in clinical practice.

Form Factor Preferences

The second question posed at the start of this work was: "What are the preferences for iPhone® and the iPad® form factor in clinical use?" No association was found between

screen sizes and ease-of-use or ease-of-navigation of the app itself. Participants rated both the same in Likert responses. The navigation bar on the iPad® did not make a significant difference to ease of use any more than the navigation popup did on the iPhone®. Participants used both controls and tended to make use of the ‘close’ button to restart navigation from the patient list.

When asked about the ease-of-use of the smaller and larger devices, the larger iPad® is preferred. Participants also preferred the ease-of-use of the app on the iPad®.

Participants made comments that the larger screen size allowed for the display of more data-tables and allowed paper-style reports to be read in a single view with no scrolling or panning. As a result, there is less need for navigation and a greater sense that the information they needed was in the Dashboard section negating the need to hunt for the correct view.

Participants also expressed concern about the risk of missing important data on the smaller screen and one participant, who preferred the information capacity of the larger screen, remained concerned about the inability of either devices’ screen to contain sufficient data to accurately form an impression of the patient’s health. For them, the risk of making an error out of ignorance outweighed the usefulness of the device.

Portability Preferences

The iPad® was perceived to be more portable than a laptop computer for the purposes of staying connected while attending a conference (e.g. at a remote location.) A remote location may be rural or, an offshore work-site setting. In this situation, for example, the portability of the device would be most important

The iPad[®] is also more portable than a desktop computer during rounds. Being able to access results remotely was perceived as a positive aspect of the system particularly while outside the clinic. The results show that despite the portable nature of handheld devices, participants did not consider the iPad[®] to be particularly portable. They did not want to be ‘hauling the iPad[®] around with you’. They were not the ‘carry-the-iPad[®]’ kind of doctors. To generalize the portability: the farther away one is from the workplace, the more portable and useful the handheld devices were perceived. The closer to the clinic or workplace, the more burdensome the iPad[®] was perceived to be. The iPad[®]’s perceived portability seemed directly proportional to the user’s distance from a main computer; while the iPad[®] perceived as a burden seemed inversely proportional to the user’s distance from a main computer.

The iPhone[®], by contrast, was perceived as being absolutely portable. No participant commented that it was a problem carrying an iPhone[®] around. Using the iPhone[®] and the app was a problem, but carrying it around was not. Participants liked the portability most about the app on the iPhone[®]. It was the most positively commented feature of the device. Alarming, one participant commented that if they were driving and received a phone call, they could also use the app to examine test results for a patient. Social irresponsibility aside, this comment was made as a genuine reaction to the iPhone[®]’s portability. This perception that ‘the iPhone[®] is portable’ did not arise as a result of the participants’ experience during the interview session. The perception had been formed a priori. All participants carried a smartphone-sized mobile device with them and one participant used it to look up a fact during the study. Participant statements regarding the size of the iPad[®] indicate that participants liked and preferred the size of the device; but

they also expressed the opinion that it was not sufficiently portable; for this they consistently preferred the iPhone® device.

Usability Problems

The third and final question posed at the start of this work was: “What types of usability problems do users have with mobile device interfaces to an electronic medical record?”

Imaging, reports and data. The formatted PDF reports, had inconsistent page numbers. All multi-page documents displayed both page number and total page count, i.e., “Page 2 of 3.” In the case of the single-page documents; however, no page number was displayed. As a result, this participant was uncertain about the number of pages in the document, and wanted to see an indication of the total page count. The fact that a document had only a single page was made implicitly by the lack of numbering. In that case it would have been better to indicate this fact in a more explicit numbered page, i.e. ‘Page 1 of 1.’

For medical imaging reports, an inconsistency existed between the iPad® and iPhone® app. Although both iPad® and iPhone® have reports with embedded imagery, the iPad® alone has a separate section for the imaging itself, i.e., without the reports. This problem caused most participants to miss one, or the other, section when using the iPad®; particularly, when they encountered the imaging section before the reports section. This pattern led some to assume that there were no imaging reports on the iPad®.

Confusion caused by the inconsistent application of units label for numerical values, i.e., sometimes the metric system of units was used, and at other times users were shown the United States customary system of units. Phone numbers that did not agree with the

geographical locations of clinic addresses also added to the level of distraction making participants' suspension of disbelief more difficult than necessary.

Participant 8:

So, he is in fact diabetic... checks his sugars in American numbers and I don't know what they mean but they don't look too bad.

The foregoing quote is unsettling because it features a judgement that this patient's sugars 'don't look bad,' juxtaposed with the notion that the participant does not know what the values mean. Perhaps the information implied, by the fact that the numbers were not decorated with any warning symbols or colours, was enough to satisfy them that the values were within normal range. Ranges that are demonstrably inconsistent and poorly documented within the app, lead one to think that; however one looks at this, the participant would be imprudent to take action, given what they've seen.

The annotative text decorations, their colours and symbolic icons used to add meaning, flag values that were out of range or otherwise questionable, were frequently discussed. The app designers used a multimodal approach to icons, altering both their colour and the shape to convey an increasing severity to users. One participant commented that they would prefer a single shape, and to alter the colours to indicate severity. A different participant, who was colour-blind, disagreed and raised a usability problem, when they were presented with difficult to differentiate colour-coded lines.

Using a single symbol and altering its colour to encode severity appears to simplify the display of information; yet, it may disorient some users who have a colour deficiency that might prevent them from perceiving different severity levels. To these users, the colour difference may not be apparent. The combination of altering the colour and icon for

visual aids in data tables seemed to be a good choice for this app. The chart colours need to be augmented by altering the line style for better usability. Perhaps a better way to simplify the presentation would be to use a single colour to get the users' attention, and alter the symbol itself to encode the severity. Doing this, especially on very visually complex views, such as the Urinalysis table view, could preserve information and reduce the sensory-overload effect of too many attention-getting features in a user interface.

Many values indicated for the normal limits on lab results were zeros. Some of these values were expressed as a range, for example the range '0 – 0', meaning apparently, from zero to zero. The zero-value entry for a normal limit is ambiguous. Zero is a value, potentially a valid one. The user may assume that it is unlikely to be a zero-length range; but, is it zero or is it simply missing data and not a valid range. The lack of clarity caused the participants to simply ignore these values and assume that the data was missing. In the case, where data is not available, it would be better to clearly express the result as 'missing,' 'unknown,' or even by a placeholder, such as a dash, rather than a misleading empty range.

Another example of a derived value was the patient's date of birth (DOB). Participants are always shown the patient's age – derived from the DOB and the current date. Participants frequently mentioned the patient's age, but rarely mentioned their DOB. In this case, the patient's age was shown. Given that this derived value was provided, participants may be justified in querying why the app did not provide other such services. One participant commented that since the app runs on a computer, the computer should do some other derivative calculations automatically. Where the results showed raw data manually processed into derived results, the participant wanted the app to be programmed

to do this calculation automatically. The app has access to the data and could perform simple or complex calculations. One participant commented that the Anion Gap was a relevant example of a useful derived value:

Participant 6:
Blood gases ... chloride is low ... I'd need a piece of paper to calculate the anion gap. It would be useful to have that calculated automatically. Since blood gases ... if you're thinking of using it on the go ...

Derived values are perhaps more relevant than their raw antecedents. The ability to perform derivative calculations is not merely a convenience. If the app is not programmed to calculate these figures, then the user may have to use another app, a different computer, or a pencil and paper, to perform them manually. All of which increases the likelihood of an error.

Controls. At least two participants were confused by the wording and appearance of text in the Orders section of the app. The issue here is caused by skeuomorphism, the design method of making a new user interface visually resemble the actual object it is intended to replace as described in Oh et al. (2013). Some report pages in this app are PDF files of printable reports. Other pages, such as the medication screen, are visually similar to paper forms. On a printed paper document, the word 'Verified' at the bottom would never be confused with a button. In an app, on the other hand, when the word 'Verified' appears at the bottom of the screen with other similar words, the presentation signifies something different: a control. The app is intended to replace paper forms in this skeuomorphic design. This intent is imparted by causing the medications screen to appear similar to a paper form. This user interface design resulted in a mixture of intents,

i.e., the app as a static source of information, versus the app as a platform for performing actions. This ambiguity is a source of confusion and created problems for users who thought the items on the form were controls when in fact they were not controls.

Navigation. Even though the navigation bar was present on the iPad®, there were 12 Navigation errors versus 13 on the iPhone®. The errors were not made when participants interacted with the iPad®'s navigation bar, instead, they were caused by an incorrect interaction with the close button. The user's reliance on a more familiar concept of 'back' button caused them to use it even though it was labelled 'close' and did not have the direct effect of sending them back to the previous view. Whatever their intentions, it did send them back to the list of patients and upon re-clicking the patient they returned to the starting point, the dashboard. This awkward procedure became a workaround to thinking about and using the navigation bar. The concept of 'take me back' could be likened to traversing a tree-like hierarchical structure; for example, from a leaf node back to the branch it came from. The navigation bar does not suggest a hierarchy. Rather, the navigation bar represents a list of all the branches and allows a user to jump from leaf of one branch to the main node of a different branch. In other words, it flattens the structure and resulted in a rejection by the participants. The participants working around the problem and restarting navigation at the bottom of the tree was preferred over the more efficient jumping immediately to another section. The general lack of familiarity with the app, no participants had used it before the session, may have contributed to this behaviour leading them to prefer traversing the tree as a way of learning its structure. Though that may be part of the reason, this researcher suspects that familiarity of years of back-button use has ingrained the back-button-seeking behaviour in participants who, in a novel

situation, stick with what they know rather than explore new means of navigation on an unfamiliar software app.

The one control that could potentially make a difference to the participants' navigation experience would be a new control that would act as the more familiar 'back' button. If this control existed, users would use it to return from whence they came up to the point where they first entered the patient record. The navigation bar/popup control was used by participants to enter new areas of a patient record; but, it was rarely used to return to the starting location. Adding a 'go-back' feature would improve the users' experience, avoid potential navigation errors, and improve the app's safety.

Chapter 7: Conclusion

This chapter concludes the discussion of research into the effects of mobile device screen sizes on clinical users' perceptions of medical application usability and safety when interfacing with critical patient information. Two popular mobile devices were examined: the iPhone[®], a smartphone-sized mobile device referred to in this chapter as “smartphone”, and the iPad[®], a larger tablet-sized computer the size of a small magazine referred to in this chapter as “tablet”. The following conclusions are presented within the context of the literature review found in Chapter 2 of this study. The contribution made by this research to health informatics knowledge and health informatics education is summarized and includes conclusions which relate to the use of mobile devices in health professional education.

Findings

Is there a correlation between the size of the handheld device's display and the perceived safety of the iPhone[®] and the iPad[®]? There is an association between screen size and the perceived safety of the handheld device. The device with the larger screen is perceived to be safer than the device with the smaller screen. This finding is consistent with similar findings found in a study measuring university student participants' actual error rates on different screen sizes (Restyandito, 2017).

What are the preferences for the iPhone[®] and the iPad[®] form factor in clinical use? Participants preferred the iPad[®] in clinical practice because of its larger size, not because they thought it was safer or easier to use. The iPhone[®] was preferred for its portability. The perceived usefulness of the iPhone[®] increased with greater imagined distance from the point of care. The preference for the larger screen is consistent with

findings from a paper by Maniar (2007) that investigated the efficiency and effectiveness of student participants information processing, i.e., learning, improved with larger handheld display sizes.

What types of usability problems do users have with the smartphone and tablet interfaces to an electronic medical record?

Consistency. Participants had problems with inconsistent page numbers, units of measurement, reference ranges and navigation modes. The participants also had difficulty interpreting the quantity and meaning of warning icons and various colours. The most consistent problem was the lack of a 'go-back' navigational button. All participants independently chose to rely on a potentially unsafe workaround.

Contributions of this Research to the Body of Knowledge

Mobile handheld general-purpose, touch-screen computers used by physicians are no longer a novelty after ten years of availability in the marketplace. The idea that mobile devices are ubiquitous is supported by evidence from this research which showed that all participants have used the devices, all own a version of the device with the smaller screen, and all were capable of completing their tasks in the experiment. No participant self-identified as a 'IT expert,' yet all were competent technology users. This study revealed that the app presented a number of usability issues, notably inconsistent controls and modes of navigation that caused frustration and error during use by participants.

Apps with usability problems result in a poor user experience. Usability problems may also cause errors and the potential for harm causes the user to lose trust in the app. An untrustworthy system will not be used and is therefore not a useful tool.

Both large and small devices, are designed to be mobile, highly-portable devices. Paradoxically, the usefulness of the device with the larger screen is adversely affected by the relative lack of portability. Participants did not want to “haul the tablet around”. If a tablet device is not perceived to be portable, then the device is not likely to be used as frequently as the more portable device, i.e., the smartphone that all participants carried with them.

Participants indicated that the tablet-sized device was safer. The tablet seemed better suited to the task of displaying the many interconnected data which made up a patient’s electronic health record. One participant expressed a desire to make use of the still larger “super tablet” for reviewing charts. Given that previous research by Jakobsen and Hornbaek (2013) found that human performance did not always improve with increased screen size, it is possible that using a “super tablet” could lead to an increase in errors and reduce safety. There is likely an optimum size for a reviewing patient charts. It may also be likely that there is a different optimum size for reviewing radiological studies. According to the literature review and results of this experiment this “sweet spot” is not yet known for devices in healthcare.

Contributions of this Research to Health Informatics Education.

Medicine is practiced with the aid of technology. Physicians use mobile handheld devices in many ways, such as: a reference, a scheduler, for calculations and as the primary source of patients’ medical history, lab results, orders, and documentation. Novel technology becomes commonplace as we enter the “smart era.” New conversations might occur near water coolers in hospital hallways: “Oh! Have you tried the new release for the prostate exam simulator?” (Farr, 2015), or, “I’ll be back in a

minute, my stethoscope needs a firmware update!” (Mazalan, 2012). The concepts of usability and their study is currently part of the Health Informatics curriculum⁴. Based on anecdotal accounts, the study of Health Informatics software usability is not part of medical school curricula. Usability is a complex subject. Poor usability can lead to decreased app-device usefulness. Being informed about an important aspect of the electronic tools used in medicine could empower physicians with a framework for tool-use and selection. The subject of software usability should be introduced in medical schools.

Software usability testing is expensive and resource intensive. For this small-scale study, 35 letters of invitation to participate were sent to prospects. Only two resulted in prompt responses. Follow-up communication secured the remaining seven participants. It took months of effort to secure a small number of reluctant candidates who have busy medical practices and many other uses for their time. Many of the studies in the literature review (see Chapter 2) were not related to health informatics research; yet, their results, (i.e., that screen size affects human performance), agree with the findings of this research which was related to health informatics. This leads the researcher to conclude that physicians may not have been needed to answer some research questions in the area of mobile device usability. In the case of this app, and these particular questions, however, physicians were needed because they were the subject-matter experts. It is likely that more efficient use of time could be made by using on-campus students, particularly medical and nursing students. Ideally, should usability be taught in medical school, a

⁴ The University of Victoria teaches usability as part of the Health Informatics curriculum as of this writing.

synergy may be attained through collaboration at universities where both health informatics and medicine are taught. School of Medicine students could be encouraged to participate in studies initiated by undergraduate or postgraduate students in the School of Health Informatics.

Alternatively, complex questions of future usability studies could be compartmentalized into: (1) those parts that can be answered using student participants, and (2) those parts that must use physician participants. For example, consider the case of a prospective study question about human performance as affected by a tablet with a stylus, or a tablet without a stylus. The researcher might attempt to gather data to answer this question by utilizing a body of “inexpensive” student participants, as the investigation would not require medical subject-matter experts. For other cases, whereby some medical subject-matter expertise is required to participate, then perhaps medical students could be considered as participants. The rationale being that although there are fewer medical than other students, they are more accessible than physicians. For cases where medical practice experience is required, only then would the researcher recruit the most “expensive” Health Informatics usability testing resource: the physician.

Software usability testing of mobile devices is complicated by the increasing numbers of devices, sizes and styles. This research demonstrates that the usability of an app is affected by the device upon which it is used. To test an app thoroughly, each device-app pair should be considered for inclusion in the research. If participants are considered a notional ‘cost’ in a study, then managing that cost by structuring questions carefully to optimize participants in an economic way, may facilitate more usability studies, or greater numbers or both.

Which is the most useful device? The smaller smartphone device is more useful than the tablet. It is more useful because: (1) it can perform the same tasks as the larger tablet device and (2) it is more portable. A physician is more likely to carry a smartphone. At any given moment, the most useful computer is the one he or she carries; therefore, the most generally *useful* device would be the smartphone.

Which size device is the safest for use with patients? The tablet was better suited to displaying patient health data. Its use resulted in fewer errors, a better layout and was preferred by all participants over the smartphone. The evidence suggests the tablet computer is the safer and more *usable* device to use with patients.

Which size screen should be procured for use in healthcare? The health care organization should choose the safer tablet device. Procuring smartphones should be left to the individual physician.

Limitations

The applicability of mobile device usability results has a major limitation: rapid obsolescence. At the beginning of this research, the iPad® had just been released to market. The iPhone® had arrived to market earlier. The app was available for both devices by the time the present reviews were completed. By the end of this research, there were a total of five iPhone® models all of different sizes and three iPad® model also with differing sizes. In Apple®'s product line there are nine devices, or nine device-app pairs to potentially evaluate. Other manufacturers also have similar-sized products adding yet more devices to test, all of which differ in many subtle ways such as user interface, input methods, sensors and screen sizes. Versions of the app used in this study

are available all of these devices (VH Chart, 2018). The app itself had also been iteratively developed with subtle UI changes. The usability assessment can be made obsolete rapidly by the introduction of new devices and their software updates.

Future Research

This study has shown that differences between two sizes of device from the same manufacturer do have different results. Usability testing is not a test of a single application software package. Rather, it is an evaluation that attempts to assess a system of hardware, operating system software, application software and human-computer interaction. The researcher is not suggesting the need to test every device-pair combination. This task, would be feasible with a team dedicated to repeatedly testing apps in use on each major release of hardware, operating system software, and application software in order to be thorough and avoid obsolescence.

The health informatics community needs a more expedient way to accomplish rapid usability testing because the Low-cost Rapid Usability Engineering (LCRUE) method described in Kushniruk and Borycki, (2006) may not be able to keep up with the accumulation of available mobile computing devices. Parts of the LCRUE method may benefit from automation, (such as those used by software developers to perform functional testing), and applying advances in other areas of technology such as: automatic pattern recognition, data-logging, and artificial intelligence. Earlier in this chapter, the cost of usability testing using ‘expensive’ participants was discussed. There are too many app-device combinations to test rapidly enough to maintain a valid set of tests for all available devices and app versions. Realistically the usability testers within the health informatics community will merely amass a set of “Rapidly Obsolete Usability Tests”

because new devices will potentially invalidate existing results faster than they can be made.

Using new software functional testing methods and novel application of software automation to solve the “Rapidly Obsolete Usability Tests” problem is akin to fighting fire with fire. The rapid advance of computing methods needs to be incorporated into LCRUE in order to measure and monitor device-app pairs for safety in healthcare. Could this testing methodology be combined with heuristic evaluation and automated into “ALCRUE”, or: Automated Low-cost Rapid Usability Engineering?

References

- 8 Years of the iPhone®: An Interactive Timeline. (2014, June 27). 8 Years of the iPhone®: An Interactive Timeline. *Time*. Retrieved from <http://time.com/2934526/apple-iphone-timeline/>
- Akiyama, M., Koshio, A., & Kaihotsu, N. (2010). Analysis of data captured by barcode medication administration system using a PDA; aiming at reducing medication errors at point of care in Japanese Red Cross Kochi Hospital. *Studies in Health Technology and Informatics*, 160(Pt 1), 774–778.
- Al-Showarah, S., AL-Jawad, N., & Sellahewa, H. (2014). Effects of User Age on Smartphone and Tablet Use, Measured with an Eye-Tracker via Fixation Duration, Scan-Path Duration, and Saccades Proportion. In *Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge* (Vol. 8514, pp. 3–14). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-07440-5_1
- Alexander, G., & Staggers, N. (2009). A systematic review of the designs of clinical technology: findings and recommendations for future research. *ANS Advances in Nursing Science*, 32(3), 252–279. <http://doi.org/10.1097/ANS.0b013e3181b0d737>
- Alsos, O. A. (2008). Attention and usability issues in mobile health information systems at point-of-care. *Studies in Health Technology and Informatics*, 136, 877–878.
- Alsos, O. A., Das, A., & Svanæs, D. (2012). Mobile health IT: the effect of user interface and form factor on doctor-patient communication., *International Journal of Medical Informatics*, 81(1), 12–28. <http://doi.org/10.1016/j.ijmedinf.2011.09.004>
- Arif, A., & Stuerzlinger, W. (2010). Predicting the cost of error correction in character-based text entry technologies. *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, 5–14.
- Auld, A. F., Wambua, N., Onyango, J., Marston, B., Namulanda, G., Ackers, M., et al. (2010). Piloting the use of personal digital assistants for tuberculosis and human immunodeficiency virus surveillance, Kenya, 2007. *The International Journal of Tuberculosis and Lung Disease : the Official Journal of the International Union Against Tuberculosis and Lung Disease*, 14(9), 1140–1146.
- Blaya, J. A., Cohen, T., Rodríguez, P., Kim, J., & Fraser, H. S. F. (2009). Personal digital assistants to collect tuberculosis bacteriology data in Peru reduce delays, errors, and workload, and are acceptable to users: cluster randomized controlled trial. *International Journal of Infectious Diseases*, 13(3), 410–418. <http://doi.org/10.1016/j.ijid.2008.09.015>
- Borycki, E. M. (2005). Identifying and preventing technology-induced error using simulations: application of usability engineering techniques. *Healthcare Quarterly*, 8(Special), 99–105.

- Borycki, E. M., & Kushniruk, A. W. (2009). Scenario-based testing of health information systems (HIS) in electronic and hybrid environments. *Studies in Health Technology and Informatics*, 143, 284–289.
- Borycki, E. M., Kushniruk, A. W., Keay, E., & Kuo, A. (2009a). A framework for diagnosing and identifying where technology-induced errors come from. *Studies in Health Technology and Informatics*, 148, 181.
- Borycki, E. M., Kushniruk, A. W., Keay, E., Nicoll, J., Anderson, J. G., & Anderson, M. M. (2009b). Toward an Integrated Simulation Approach for Predicting and Preventing Technology-Induced Errors in Healthcare: Implications for Healthcare Decision-Makers. *Healthcare Quarterly*, 12(Spec. 2009), 90–98.
- Brooke, J. (2013). System Usability Scale (SUS). Retrieved from <http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>
- Büchel, D., Baumann, T., & Matern, U. (2007). Usability of radio-frequency devices in surgery. *HCI and Usability for Medicine and Health Care*, 97–104.
- Carless, M. (2004, January 1). Screenplay Format for TV Shows. Retrieved June 5, 2018, from <http://downloads.bbc.co.uk/writersroom/scripts/screenplaytv.pdf>
- Carroll, A. E., DiMeglio, L. A., Stein, S., & Marrero, D. G. (2011). Using a cell phone-based glucose monitoring system for adolescent diabetes management. *The Diabetes Educator*, 37(1), 59–66. <http://doi.org/10.1177/0145721710387163>
- Carvalho, C. J., Borycki, E. M., & Kushniruk, A. W. (2009). Using heuristic evaluations to assess the safety of health information systems. *Studies in Health Technology and Informatics*, 143, 297–301.
- Catalyst. (n.d.). Smartphone Behaviour in Canada and the Implications for Marketers in 2016. Retrieved December 24, 2017, from <http://catalyst.ca/2016-canadian-smartphone-behaviour/>
- Chen, C.-M., & Lin, Y.-J. (2014). Effects of different text display types on reading comprehension, sustained attention and cognitive load in mobile reading contexts. *Interactive Learning Environments*, 1–19. <http://doi.org/10.1080/10494820.2014.891526>
- Cimino, J. J., Patel, V. L., & Kushniruk, A. W. (2002). The patient clinical information system (PatCIS): technical solutions for and experience with giving patients access to their electronic medical records, 68(1-3), 113–127.
- Coiera, E., Ash, J., & Berg, M. (2016). The Unintended Consequences of Health Information Technology Revisited. *Yearbook of Medical Informatics*, (1), 163–169. <http://doi.org/10.15265/IY-2016-014>

- Coopmans, V. C., & Biddle, C. (2008). CRNA performance using a handheld, computerized, decision-making aid during critical events in a simulated environment: a methodologic inquiry. *AANA Journal*, 76(1), 29–35.
- Cunningham, T., Geller, E., & Clarke, S. (2008). Impact of electronic prescribing in a hospital setting: A process-focused evaluation, 77(8), 546–554.
- Currie, L. M. (2005). Evaluation frameworks for nursing informatics, 74(11-12), 908–916. <http://doi.org/10.1016/j.ijmedinf.2005.07.007>
- Dabbs, A. D., MYERS, B. A., MC CURRY, K. R., DUNBAR-JACOB, J., HAWKINS, R. P., BEGEY, A., & DEW, M. A. (2009). User-Centered Design and Interactive Health Technologies for Patients. *CIN: Computers, Informatics, Nursing*, 27(3), 175–183. <http://doi.org/10.1097/NCN.0b013e31819f7c7c>
- Dallenbach, M. F., Bovier, P. A., & Desmeules, J. (2007). Detecting drug interactions using personal digital assistants in an out-patient clinic. *Qjm*, 100(11), 691–697. <http://doi.org/10.1093/qjmed/hcm088>
- Dawson, J., & Kushniruk, A. W. (2009). Improving the efficiency and accuracy of a tablet PC interface for computerized provider order entry through usability evaluation and provision of data entry strategies. *Studies in Health Technology and Informatics*, 143, 447–452.
- Denham, C., Classen, D., Swenson, S., Henderson, M., Zeltner, T., Bates, D. (2013), Safe use of electronic health records and health information technology systems: trust but verify. *Journal of Patient Safety*, 9(4), 177–189
- Devine, E. B., Hansen, R. N., Wilson-Norton, J. L., Lawless, N. M., Fisk, A. W., Blough, D. K., et al. (2010). The impact of computerized provider order entry on medication errors in a multispecialty group practice. *Journal of the American Medical Informatics Association*, 17(1), 78–84. <http://doi.org/10.1197/jamia.M3285>
- Dey, A., Jarvis, G., Sandor, C., & Reitmayr, G. (2012). Tablet versus phone: Depth perception in handheld augmented reality (pp. 187–196). Presented at the ISMAR '12: Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR, IEEE Computer Society. <http://doi.org/10.1109/ISMAR.2012.6402556>
- Dolan, B. (2009). Dolan: FDA may regulate iPhone® Health Apps. Retrieved from <http://bit.ly/u2nna>
- Dollarhide, A. W., Rutledge, T., Weinger, M. B., & Dresselhaus, T. R. (2008). Use of a Handheld Computer Application for Voluntary Medication Event Reporting by Inpatient Nurses and Physicians. *Journal of General Internal Medicine*, 23(4), 418–422. <http://doi.org/10.1007/s11606-007-0404-0>

- Eirik Årsand, N. T. G. Ø. G. H. (2010). Mobile Phone-Based Self-Management Tools for Type 2 Diabetes: The Few Touch Application. *Journal of Diabetes Science and Technology*, 4(2), 328.
- Epstein, D. H., & Preston, K. L. (2010). Daily life hour by hour, with and without cocaine: an ecological momentary assessment study. *Psychopharmacology*, 211(2), 223–232. <http://doi.org/10.1007/s00213-010-1884-x>
- Farr, C. (2015, May 26). Robotic Butt Helps Medical Students Learn Professional Intimacy. Retrieved August 5, 2018, from <https://www.kqed.org/futureofyou/3339/robotic-butt-helps-medical-students-learn-professional-intimacy>
- Findlater, L., & McGrenere, J. (2008). Impact of screen size on performance, awareness, and user satisfaction with adaptive graphical user interfaces. Presented at the Proceedings of the SIGCHI Conference on
- Fritz, F., Balhorn, S., Riek, M., Breil, B., & Dugas, M. (2012). Qualitative and quantitative evaluation of EHR-integrated mobile patient questionnaires regarding usability and cost-efficiency. <http://doi.org/10.1016/j.ijmedinf.2011.12.008>
- Fromme, E. K., Kenworthy-Heinige, T., & Hribar, M. (2010). Developing an easy-to-use tablet computer application for assessing patient-reported outcomes in patients with cancer. *Supportive Care in Cancer*, 19(6), 815–822. <http://doi.org/10.1007/s00520-010-0905-y>
- Gallagher, J. M., Stewart, T. V., Pathak, P. K., Werner, J. J., Dickinson, L. M., & Hickner, J. M. (2008). Data Collection Outcomes Comparing Paper Forms With PDA Forms in an Office-Based Patient Survey. *The Annals of Family Medicine*, 6(2), 154–160. <http://doi.org/10.1370/afm.762>
- García-Sáez, G., Hernando, M. E., Martínez-Sarriegui, I., Rigla, M., Torralba, V., Brugués, E., et al. (2009). Architecture of a wireless Personal Assistant for telemedical diabetes care, 78(6), 391–403. <http://doi.org/10.1016/j.ijmedinf.2008.12.003>
- Goodman, K. W., & Miller, R. A. (2006). 10 Ethics and Health Informatics: Users, Standards, and Outcomes. In *Biomedical Informatics* (p. 1037). Springer Verlag.
- Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? *Psychological Bulletin*, 83(2), 314–320. <http://doi.org/10.1037/0033-2909.83.2.314>
- Hachizuka, M., Yoshiuchi, K., Yamamoto, Y., Iwase, S., Nakagawa, K., Kawagoe, K., & Akabayashi, A. (2010). Development of a Personal Digital Assistant (PDA) System To Collect Symptom Information from Home Hospice Patients. *Journal of Palliative Medicine*, 13(6), 647–651. <http://doi.org/10.1089/jpm.2009.0350>

Haller, G., Haller, D. M., Courvoisier, D. S., & Lovis, C. (2009). Handheld vs. Laptop Computers for Electronic Data Collection in Clinical Research: A Crossover Randomized Trial. *Journal of the American Medical Informatics Association*, *16*(5), 651–659. <http://doi.org/10.1197/jamia.M3041>

Hammond, C. E., & Nixon, M. A. (2011). The reliability of a handheld wound measurement and documentation device in clinical practice. *Journal of Wound, Ostomy, and Continence Nursing : Official Publication of the Wound, Ostomy and Continence Nurses Society / WOCN*, *38*(3), 260–264. <http://doi.org/10.1097/WON.0b013e318215fc60>

Hayden, R. T., Patterson, D. J., Jay, D. W., Cross, C., Dotson, P., Possel, R. E., et al. (2008). Computer-assisted bar-coding system significantly reduces clinical laboratory specimen identification errors in a pediatric oncology hospital. *The Journal of Pediatrics*, *152*(2), 219–224. <http://doi.org/10.1016/j.jpeds.2007.08.021>

Hess, R., Santucci, A., McTigue, K., Fischer, G., & Kapoor, W. (2008). Patient Difficulty Using Tablet Computers to Screen in Primary Care. *Journal of General Internal Medicine*, *23*(4), 476–480. <http://doi.org/10.1007/s11606-007-0500-1>

Horsky, J., Kuperman, G. J., & Patel, V. L. (2005). Comprehensive analysis of a medication dosing error related to CPOE. *Journal of the American Medical Informatics Association : JAMIA*, *12*(4), 377–382. <http://doi.org/10.1197/jamia.M1740>

Hysong, S. J., Sawhney, M., Wilson, L., Sittig, D. F., Esquivel, A., Watford, M., et al. (2009). Improving outpatient safety through effective electronic communication: a study protocol. *Implementation Science*, *4*(62).

iOS Human Interface Guidelines. (2012). iOS Human Interface Guidelines (pp. 1–184). Apple Computer.

iOS version history. (n.d.). iOS version history. Retrieved February 28, 2016, from https://en.wikipedia.org/wiki/IOS_version_history

Isaacson, W. (2011). Steve Jobs. Simon and Schuster.

Jakobsen, M. R., & Hornbæk, K. (2013). Interactive visualizations on large and small displays: the interrelation of display size, information space, and scale. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2336–2345. <http://doi.org/10.1109/TVCG.2013.170>

Jayawardena, S., Eisdorfer, J., Indulkar, S., Pal, S. A., Sooriabalan, D., & Cucco, R. (2007). Prescription errors and the impact of computerized prescription order entry system in a community-based hospital. *American Journal of Therapeutics*, *14*(4), 336–340. <http://doi.org/10.1097/01.mjt.0000209681.22077.b9>

- Johansson, P. E., Petersson, G. I., & Nilsson, G. C. (2010). Personal digital assistant with a barcode reader—A medical decision support system for nurses in home care, *79*(4), 232–242. <http://doi.org/10.1016/j.ijmedinf.2010.01.004>
- Johansson, P., Petersson, G., & Nilsson, G. (2011). Experience of using a personal digital assistant in nursing practice - a single case study. *Journal of Nursing Management*, *19*(7), 855–862. <http://doi.org/10.1111/j.1365-2834.2011.01261.x>
- John, R., Buschman, P., Chaszar, M., Honig, J., Mendonca, E., & Bakken, S. (2007). Development and evaluation of a PDA-based decision support system for pediatric depression screening. *Studies in Health Technology and Informatics*, *129*(Pt 2), 1382–1386.
- Joundi, R. A., Brittain, J.-S., Jenkinson, N., Green, A. L., & Aziz, T. (2011). Rapid tremor frequency assessment with the iPhone® accelerometer. *Parkinsonism & Related Disorders*, *17*(4), 288–290. <http://doi.org/10.1016/j.parkreldis.2011.01.001>
- Kanstrup, A. M., & Stage, J. (2009). From Paper to PDA: Design and Evaluation of a Clinical Ward Instruction on a Mobile Device. In *Human-Computer Interaction – INTERACT 2009* (Vol. 5727, pp. 670–683). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-03658-3_72
- Karahoca, A., Bayraktar, E., Tatoglu, E., & Karahoca, D. (2010). Information system design for a hospital emergency department: A usability analysis of software prototypes. *Journal of Biomedical Informatics*, *43*(2), 224–232. <http://doi.org/10.1016/j.jbi.2009.09.002>
- Kastner, M., Li, J., Lottridge, D., Marquez, C., Newton, D., & Straus, S. E. (2010). Development of a prototype clinical decision support tool for osteoporosis disease management: a qualitative study of focus groups. *BMC Medical Informatics and Decision Making*, *10*(1), 40. <http://doi.org/10.1186/1472-6947-10-40>
- Khajouei, R., de Jongh, D., & Jaspers, M. W. M. (2009). Usability evaluation of a computerized physician order entry for medication ordering. *Studies in Health Technology and Informatics*, *150*, 532–536.
- Kim, Dong Keun, Kim, E. Y., Yang, K. H., Lee, C. K., & Yoo, S. K. (2011). A mobile tele-radiology imaging system with JPEG2000 for an emergency care. *Journal of Digital Imaging*, *24*(4), 709–718. <http://doi.org/10.1007/s10278-010-9335-0>
- Kim, Junetae, Lee, Y., Lim, S., Kim, J. H., Lee, B., & Lee, J.-H. (2017). What Clinical Information Is Valuable to Doctors Using Mobile Electronic Medical Records and When? *Journal of Medical Internet Research*, *19*(10), e340. <http://doi.org/10.2196/jmir.8128>
- Kuiper, R. (2008). Use of Personal Digital Assistants to Support Clinical Reasoning in Undergraduate Baccalaureate Nursing Students. *CIN: Computers, Informatics, Nursing*, *26*(2), 90–98. <http://doi.org/10.1097/01.NCN.0000304776.40531.bc>

- Kushniruk, A. W., & Borycki, E. M. (2006). Low-cost rapid usability engineering: designing and customizing usable healthcare information systems. *Healthcare Quarterly (Toronto, Ont.)*, 9(4), 98–100–102.
- Kushniruk, A. W., & Borycki, E. M. (2015). Development of a Video Coding Scheme for Analyzing the Usability and Usefulness of Health Information Systems. *Studies in Health Technology and Informatics*, 218, 40599.
- Kushniruk, A. W., Borycki, E. M., Anderson, J. G., & Anderson, M. M. (2009). Preventing technology-induced errors in healthcare: the role of simulation. *Studies in Health Technology and Informatics*, 143, 273–276.
- Kushniruk, A. W., Nohr, C., & Borycki, E. M. (2016). Human Factors for More Usable and Safer Health Information Technology: Where Are We Now and Where do We Go from Here? *Yearbook of Medical Informatics*, (1), 120–125. <http://doi.org/10.15265/IY-2016-024>
- Kushniruk, A. W., Patel, V. L., & Cimino, J. J. (1997). Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces. *Proceedings : a Conference of the American Medical Informatics Association. AMIA Fall Symposium*, 218–222.
- Kushniruk, A. W., Triola, M. M., Borycki, E. M., Stein, B., & Kannry, J. L. (2005a). Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. *International Journal of Medical Informatics*, 74(7-8), 519–526. <http://doi.org/10.1016/j.ijmedinf.2005.01.003>
- Kushniruk, A. W., Triola, M., & Borycki, E. M. (2005b). Technology induced error and usability: The relationship between usability problems and prescription errors when using a handheld application, (74), 519–526.
- Lee, N.-J., Chen, E. S., Currie, L. M., Donovan, M., Hall, E. K., Jia, H., et al. (2009). The effect of a mobile clinical decision support system on the diagnosis of obesity and overweight in acute and primary care encounters. *ANS Advances in Nursing Science*, 32(3), 211–221. <http://doi.org/10.1097/ANS.0b013e3181b0d6bf>
- Lilholt, L. H., Pedersen, S. S., Madsen, I., Nielsen, P. H., Boye, N., Andersen, S. K., & Nøhr, C. (2006). Development of methods for usability evaluations of EHR systems. *Studies in Health Technology and Informatics*, 124, 341–346.
- Luckmann, R., & Vidal, A. (2010). Design of a handheld electronic pain, treatment and activity diary. *Journal of Biomedical Informatics*, 43(5 Suppl), S32–6. <http://doi.org/10.1016/j.jbi.2010.05.005>
- Mahlmeister, L. (2010). Human Factors and Error in Perinatal Care: The Interplay Between Nurses, Machines, and the Work Environment. *The Journal of Perinatal & Neonatal Nursing*, 24(1), 12.

- Malone, D. C., & Saverno, K. R. (2012). Evaluation of a wireless handheld medication management device in the prevention of drug-drug interactions in a Medicaid population. *Journal of Managed Care Pharmacy : JMCP*, 18(1), 33–45.
- Maniar, N. (2007). M-learning to teach university students. *World Conference on Educational Multimedia*.
- Marceau, L. D., Link, C., Jamison, R. N., & Carolan, S. (2007). Electronic Diaries as a Tool to Improve Pain Management: Is There Any Evidence? *Pain Medicine*, 8(s3 Computer and), S101–S109. <http://doi.org/10.1111/j.1526-4637.2007.00374.x>
- Martin, L., Ariza, A., Thomson, J., & Binns, H. (2008). Seconds for Care: Evaluation of Five Health Supervision Visit Topics Using a New Method. *The Journal of Pediatrics*, 153(5), 706–711.e2. <http://doi.org/10.1016/j.jpeds.2008.05.001>
- Mazalan, S. M. (2012). An Electronic Stethoscope Based on Wireless Bluetooth.
- Matts, J. P., & Lachin, J. M. (1988). Properties of permuted-block randomization in clinical trials. *Controlled Clinical Trials*, 9(4), 327–344. [http://doi.org/10.1016/0197-2456\(88\)90047-5](http://doi.org/10.1016/0197-2456(88)90047-5)
- Miller, A., Pilcher, D., Mercaldo, N., Leong, T., Scheinkestel, C., & Schildcrout, J. (2010). What can paper-based clinical information systems tell us about the design of computerized clinical information systems (CIS) in the ICU? *Australian Critical Care : Official Journal of the Confederation of Australian Critical Care Nurses*, 23(3), 130–140. <http://doi.org/10.1016/j.aucc.2010.02.001>
- Miller, D. P., Weaver, K. E., Case, L. D., Babcock, D., Lawler, D., Denizard-Thompson, N., et al. (2017). Usability of a Novel Mobile Health iPad® App by Vulnerable Populations. *JMIR mHealth and uHealth*, 5(4), e43. <http://doi.org/10.2196/mhealth.7268>
- Momtahan, K. L., Burns, C. M., Sherrard, H., Mesana, T., & Labinaz, M. (2007). Using personal digital assistants and patient care algorithms to improve access to cardiac care best practices. *Studies in Health Technology and Informatics*, 129(Pt 1), 117–121.
- Montague, E. N., Kleiner, B. M., & Iii, W. W. W. (2009). Empirically understanding trust in medical technology. *International Journal of Industrial Ergonomics*, 39(4), 628–634. <http://doi.org/10.1016/j.ergon.2009.01.004>
- Murphy, K., Wong, F., & Martin, L. (2009). Ongoing evaluation of ease-of-use... [Stud Health Technol Inform. 2009] - PubMed - NCBI. *Studies in Health*
- Ninos, K., Spiros, K., Glotsos, D., Georgiadis, P., Sidiropoulos, K., Dimitropoulos, N., et al. (2010). Development and evaluation of a PDA-based teleradiology terminal in thyroid nodule diagnosis. *Journal of Telemedicine and Telecare*, 16(5), 232–236. <http://doi.org/10.1258/jtt.2010.090512>

Nolan, M. (2008). Remote CPOE error - a situation that is more than remotely possible. *EJHP Practice*, 14(4), 66–67.

Oh, In-Kyun and Jung, Suk-Gill (2013). A Study on Skeuomorphism Characteristics in Application User Interface Design of Apple. *Journal of Digital Design*, 13(2), 463- 472.

PalmOne. (n.d.). PalmOne. Retrieved January 10, 2018, from <https://web.archive.org/web/20040805122321/http://www.palmone.com/us/company/corporate/timeline.html>

Park, J., Tymitz, K., Engel, A. M., & Welling, R. E. (2007). Computerized rounding in a community hospital surgery residency program. *Journal of Surgical Education*, 64(6), 357–360. <http://doi.org/10.1016/j.jsurg.2007.04.001>

Pirnejad, H., Niazkhani, Z., van der Sijs, H., Berg, M., & Bal, R. (2009). Evaluation of the impact of a CPOE system on nurse-physician communication--a mixed method study. *Methods of Information in Medicine*, 48(4), 350–360. <http://doi.org/10.3414/ME0572>

Quach, B., & Balakrishnan, M. (2009). Implementation of integrated wireless health monitoring network. *Proceedings of the Ist*

Ra, M., Muhammad, M.S., Lim, C., Han, S., Jung, C. and Kim, W., "Smartphone-Based Point-of-Care Urinalysis Under Variable Illumination." *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 6, pp. 1-11, 2018, doi: 10.1109/JTEHM.2017.2765631

Ramey, J., Fung, K. M., & Hassell, L. A. (2011). Use of mobile high-resolution device for remote frozen section evaluation of whole slide images. *Journal of Pathology Informatics*, 2, 41. <http://doi.org/10.4103/2153-3539.84276>

Raptis, D., Tselios, N., Kjeldskov, J., & Skov, M. B. (2013). Does size matter?: investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency. *the 15th international conference* (pp. 127–136). New York, New York, USA: ACM. <http://doi.org/10.1145/2493190.2493204>

Ray, M. N., Houston, T. K., Yu, F. B., Menachemi, N., Maisiak, R. S., Allison, J. J., & Berner, E. S. (2006). Development and testing of a scale to assess physician attitudes about handheld computers with decision support. *Journal of the American Medical Informatics Association : JAMIA*, 13(5), 567–572. <http://doi.org/10.1197/jamia.M2096>

Redden, E. S., Elliott, L. R., Pettitt, R. A., & Carstens, C. B. (2009a). A tactile option to reduce robot controller size. *Journal on Multimodal User Interfaces*, 2(3-4), 205–216. <http://doi.org/10.1007/s12193-009-0019-3>

Redden, E. S., Pettitt, R. A., Carstens, C. B., & Elliott, L. R. (2008). Scalability of robotic displays: Display size Investigation.

- Redden, E. S., Pettitt, R. A., Carstens, C. B., & Elliott, L. R. (2009b). Scaling robotic displays: Visual and multimodal options for navigation by dismounted soldiers.
- Restyandito, E. K. (2017). The Effect of Screen Size and Interaction Style on Mobile Device Usability. *Engineering Letters*, 25(4), 5.
- Riazzoli, J., Nilsson, J. A., Teleman, A., Petersson, I. F., Rantapaa-Dahlqvist, S., Jacobsson, L. T. H., & van Vollenhoven, R. F. (2010). Patient-reported 28 swollen and tender joint counts accurately represent RA disease activity and can be used to assess therapy responses at the group level. *Rheumatology*, 49(11), 2098–2103. <http://doi.org/10.1093/rheumatology/keq197>
- Rivera, M. L., Donnelly, J., Parry, B. A., Dinizio, A., Johnson, C. L., Kline, J. A., & Kabrhel, C. (2008). Prospective, randomized evaluation of a personal digital assistant-based research tool in the emergency department. *BMC Medical Informatics and Decision Making*, 8(1), 3. <http://doi.org/10.1186/1472-6947-8-3>
- Robinson, T. J., DuVall, S., & Wiggins, R. (2008). Creation and Usability Testing of a Web-Based Pre-Scanning Radiology Patient Safety and History Questionnaire Set. *Journal of Digital Imaging*, 22(6), 641–647. <http://doi.org/10.1007/s10278-008-9148-6>
- Rölleke, T. (2009). Incident reports to BfArM support the importance of usability for patient safety. *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, 298–300.
- Rubin, J., Chisnell, D. (2009). Handbook of Usability Testing.
- Sadasivam, R. S., Gathibandhe, V., Tanik, M. M., & Willig, J. H. (2010). Development of a Point-of-Care HIV/Aids Medication Dosing Support System Using the Android Mobile Platform. *Journal of Medical Systems*. <http://doi.org/10.1007/s10916-010-9619-4>
- Saleem, J. J., Russ, A. L., Justice, C. F., Hagg, H., Ebright, P. R., Woodbridge, P. A., & Doebbeling, B. N. (2009). Exploring the persistence of paper with the electronic health record, 78(9), 618–628. <http://doi.org/10.1016/j.ijmedinf.2009.04.001>
- Santos, dos, I., Teixeira, D., Ferraz, F., & Carvalho, P. (2008). The use of a simulator to include human factors issues in the interface design of a nuclear power plant control room. *Journal of Loss Prevention in the Process Industries*, 21(3), 227–238.
- Sasson, J., Alvero, A., & Austin, J. (2006). Effects of Process and Human Performance Improvement Strategies. *Journal of Organizational Behavior Management*, 26(3), 43–78. http://doi.org/10.1300/J075v26n03_02
- Scandurra, I., Hagglund, M., Koch, S., & Lind, M. (2008). Usability Laboratory Test of a Novel Mobile Homecare Application with Experienced Home Help Service Staff. *The Open Medical Informatics Journal*, 2(1), 117–128. <http://doi.org/10.2174/1874431100802010117>

- Schnall, R., Velez, O., John, R. M., & Bakken, S. (2011). Psychometric evaluation of the attitudes toward Handheld Decision Support Software Scale in a sample of nursing students. *Computers, Informatics, Nursing : CIN*, 29(4), 251–255. <http://doi.org/10.1097/NCN.0b013e3181f9dc1d>
- Seebregts, C. J., Zwarenstein, M., Mathews, C., Fairall, L., Flisher, A. J., Seebregts, C., et al. (2009). Handheld computers for survey and trial data collection in resource-poor settings: Development and evaluation of PDACT, a Palm™ Pilot interviewing system, 78(11), 721–731. <http://doi.org/10.1016/j.ijmedinf.2008.10.006>
- Sevick, M. A., Zickmund, S., Korytkowski, M., Piraino, B., Sereika, S., Mihalko, S., et al. (2008). Design, feasibility, and acceptability of an intervention using personal digital assistant-based self-monitoring in managing type 2 diabetes. *Contemporary Clinical Trials*, 29(3), 396–409. <http://doi.org/10.1016/j.cct.2007.09.004>
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, 4(1), 1–32. <http://doi.org/10.1146/annurev.clinpsy.3.022806.091415>
- Shively, M., Rutledge, T., Rose, B. A., Graham, P., Long, R., Stucky, E., et al. (2011). Real-time assessment of nurse work environment and stress. *Journal for Healthcare Quality : Official Publication of the National Association for Healthcare Quality*, 33(1), 39–48. <http://doi.org/10.1111/j.1945-1474.2010.00093.x>
- Shrestha, S. (2007). Mobile web browsing: usability study. ... *Of the 4th International Conference on Mobile ...* <http://doi.org/10.1145/2467696.2467735>
- Siek, K. A., Ross, S. E., Khan, D. U., Haverhals, L. M., Cali, S. R., & Meyers, J. (2010). Colorado Care Tablet: the design of an interoperable Personal Health Application to help older adults with multimorbidity manage their medications. *Journal of Biomedical Informatics*, 43(5 Suppl), S22–6. <http://doi.org/10.1016/j.jbi.2010.05.007>
- Sowan, A. K., Mohamed, G., Soeken, K., Mills, M. E., Johantgen, M., & Vaidya, V. (2006). A comparison of medication administrations errors using CPOE orders vs. handwritten orders for pediatric continuous drug infusions. *AMIA Annual Symposium Proceedings / AMIA Symposium AMIA Symposium*, 1105.
- Spindler, M., Schuessler, M., Martsch, M., & Dachsel, R. (2014). Pinch-drag-flick vs. spatial input (pp. 1113–1122). Presented at the the 32nd annual ACM conference, New York, New York, USA: ACM Press. <http://doi.org/10.1145/2556288.2557028>
- Sprigle, S., Nemeth, M., & Gajjala, A. (2012). Iterative design and testing of a hand-held, non-contact wound measurement device. *Journal of Tissue Viability*, 21(1), 17–26. <http://doi.org/10.1016/j.jtv.2011.09.001>
- Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2003). Patient compliance with paper and electronic diaries. *Controlled Clinical Trials*, 24(2), 182–199. [http://doi.org/10.1016/S0197-2456\(02\)00320-3](http://doi.org/10.1016/S0197-2456(02)00320-3)

- Su, K.-W., & Liu, C.-L. (2010). A Mobile Nursing Information System Based on Human-Computer Interaction Design for Improving Quality of Nursing. *Journal of Medical Systems*. <http://doi.org/10.1007/s10916-010-9576-y>
- Svanæs, D., Alsos, O. A., & Dahl, Y. (2010). Usability testing of mobile ICT for clinical settings: methodological and practical challenges., *79*(4), e24–34. <http://doi.org/10.1016/j.ijmedinf.2008.06.014>
- Svanæs, D., Das, A., & Alsos, O. A. (2008). The contextual nature of usability and its relevance to medical informatics. *Studies in Health Technology and Informatics*, *136*, 541–546.
- Takeuchi, K., & Kennelly, P. J. (2010). iSeismometer: A geoscientific iPhone® application. *Computers & Geosciences*, *36*(4), 573–575. <http://doi.org/10.1016/j.cageo.2009.09.012>
- Tani, S., Marukami, T., Matsuda, A., Shindo, A., Takemoto, K., & Inada, H. (2009). Development of a Health Management Support System for Patients with Diabetes Mellitus at Home. *Journal of Medical Systems*, *34*(3), 223–228. <http://doi.org/10.1007/s10916-008-9233-x>
- Teramoto, K., Kuwata, S., Kushniruk, A. W., Borycki, E. M., Nomura, M., & Kondoh, H. (2010). Evaluation of response time delay of a pen-tablet system in thin-client computing environments: implications for usability. *Studies in Health Technology and Informatics*, *156*, 47–56.
- Thriemer, K., Ley, B., Ame, S. M., Puri, M. K., Hashim, R., Chang, N. Y., et al. (2012). Replacing paper data collection forms with electronic data entry in the field: findings from a study of community-acquired bloodstream infections in Pemba, Zanzibar. *BMC Research Notes*, *5*(1), 113. <http://doi.org/10.1186/1756-0500-5-113>
- Tseng, F.-Y., Chao, C.-J., Feng, W.-Y., & Hwang, S.-L. (2013). Effects of display modality on critical battlefield e-map search performance. *Behaviour & Information Technology*, *32*(9), 888–901. <http://doi.org/10.1080/0144929X.2012.702286>
- Turner-Bowker, D. M., Saris-Baglana, R. N., Smith, K. J., DeRosa, M. A., Paulsen, C. A., & Hogue, S. J. (2011). Heuristic evaluation and usability testing of a computerized patient-reported outcomes survey for headache sufferers. *Telemedicine Journal and E-Health : the Official Journal of the American Telemedicine Association*, *17*(1), 40–45. <http://doi.org/10.1089/tmj.2010.0114>
- Valdez, A. C., Ziefle, M., Horstmann, A., Herding, D., & Schroeder, U. (2010). Task performance in mobile and ambient interfaces. Does size matter for usability of electronic diabetes assistants? (pp. 514–521). IEEE.
- van der Sijs, H., Aarts, J., van Gelder, T., Berg, M., & Vulto, A. (2008). Turning off frequently overridden drug alerts: limited opportunities for doing it safely. *Journal of the American Medical Informatics Association*, *15*(4), 439.

- van der Sijs, H., van Gelder, T., Vulto, A., Berg, M., & Aarts, J. (2010). Understanding handling of drug safety alerts: a simulation study, *79*(5), 361–369. <http://doi.org/10.1016/j.ijmedinf.2010.01.008>
- van Duinen, M., Rickelt, J., & Griez, E. (2008). Validation of the electronic Visual Analogue Scale of Anxiety. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *32*(4), 1045–1047. <http://doi.org/10.1016/j.pnpbp.2008.02.002>
- Vawdrey, D. K., Wilcox, L. G., Collins, S. A., Bakken, S., Feiner, S., Boyer, A., & Restaino, S. W. (2011). A tablet computer application for patients to participate in their hospital care. *AMIA Annual Symposium Proceedings / AMIA Symposium AMIA Symposium, 2011*, 1428–1435.
- , August 6). VH Chart. Retrieved August 6, 2018, from <https://vitalhub.com/products-and-solutions/vh-chart/>
- Vinney, L. A., Grade, J. D., & Connor, N. P. (2012). Feasibility of using a handheld electronic device for the collection of patient reported outcomes data from children. *Journal of Communication Disorders*, *45*(1), 12–19. <http://doi.org/10.1016/j.jcomdis.2011.10.001>
- Wenzer, H. S., Böttger, U., & Boye, N. (2006). A socio-technical study of an ubiquitous CPOE-system in local use. *Studies in Health Technology and Informatics*, *124*, 326–332.
- Westbrook, J. I., & Woods, A. (2009). Development and testing of an observational method for detecting medication administration errors using information technology. *Studies in Health Technology and Informatics*, *146*, 429–433.
- Wikipedia. (n.d.). IBM Personal Computer. Retrieved May 27, 2012, from http://en.wikipedia.org/wiki/IBM_Personal_Computer
- Wilson, F. A., & Stimpson, J. P. (2010). Trends in Fatalities From Distracted Driving in the United States, 1999 to 2008. *American Journal of Public Health*, *100*(11), 2213–2219. <http://doi.org/10.2105/AJPH.2009.187179>
- Wolpin, S., Nguyen, H. Q., Donesky-Cuenco, D., Carrieri-Kohlman, V., & Doorenbos, A. (2011). Effects of automated prompts for logging symptom and exercise data on mobile devices in patients with chronic obstructive pulmonary disease. *Computers, Informatics, Nursing : CIN*, *29*(2 Suppl), TC3–8. <http://doi.org/10.1097/NCN.0b013e3182155203>
- Woods, C. A., & Cumming, B. (2009). The Impact of Test Medium on Use of Visual Analogue Scales. *Eye & Contact Lens: Science & Clinical Practice*, *35*(1), 6–10. <http://doi.org/10.1097/ICL.0b013e3181909b03>
- Wu, R. C., Orr, M. S., Chignell, M., & Straus, S. E. (2008). Usability of a mobile electronic medical record prototype: a verbal protocol analysis. *Informatics for Health & Social Care*, *33*(2), 139–149. <http://doi.org/10.1080/17538150802127223>

- Yaman, H., Yavuz, E., Er, A., Vural, R., Albayrak, Y., Yardimci, A., & Asilkan, Ö. (2015). The use of mobile smart devices and medical apps in the family practice setting. *Journal of Evaluation in Clinical Practice*, n/a–n/a. <http://doi.org/10.1111/jep.12476>
- Yap, K. Y.-L., Chui, W. K., & Chan, A. (2011). Development of a chemotherapy regimen interaction database for the mobile internet: detecting interactions with psychotropics through OncoRx-MI. *Informatics for Health & Social Care*, 36(3), 132–146. <http://doi.org/10.3109/17538157.2011.584996>
- Yee, K. C., Wong, M. C., & Turner, P. (2006). Medical error management and the role of information technology--a new approach to investigating medical handover in acute care settings. *Studies in Health Technology and Informatics*, 124, 679–684.
- Yu, P., de Courten, M., Pan, E., Galea, G., & Pryor, J. (2009). The development and evaluation of a PDA-based method for public health surveillance data collection in developing countries, 78(8), 532–542. <http://doi.org/10.1016/j.ijmedinf.2009.03.002>
- Zemrowski, K. (2008). Impacts of Increasing Reliance on Automation in Air Traffic Control Systems. *Systems Conference, 2008 2nd Annual IEEE*, 435–440.
- Ziefle, M. (2010). Information presentation in small screen devices: The trade-off between visual density and menu foresight. *Applied Ergonomics*, 41(6), 719–730. <http://doi.org/10.1016/j.apergo.2010.03.001>
- Zwarenstein, M. F., Dainty, K. N., Quan, S., Kiss, A., & Adhikari, N. K. J. (2007). A cluster randomized trial evaluating electronic prescribing in an ambulatory care setting. *Trials*, 8, 28. <http://doi.org/10.1186/1745-6215-8-28>

Appendix 1: Ethics Approval



Office of Research Services | Human Research Ethics Board
 Administrative Services Building Rm B202 PO Box 1700 STN CSC Victoria BC V8W 2Y2 Canada
 T 250-472-4545 | F 250-721-8960 | uvic.ca/research | ethics@uvic.ca

Certificate of Approval

PRINCIPAL INVESTIGATOR: Simon Minshall	ETHICS PROTOCOL NUMBER 16-275
UVic STATUS: Master's Student	Minimal Risk Review - Delegated
UVic DEPARTMENT: HEIS	ORIGINAL APPROVAL DATE: 27-Jul-16
SUPERVISOR: Dr. Andre Kushniruk	APPROVED ON: 27-Jul-16
	APPROVAL EXPIRY DATE: 26-Jul-17
PROJECT TITLE: Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records	
RESEARCH TEAM MEMBER None	
DECLARED PROJECT FUNDING: None	
CONDITIONS OF APPROVAL	
<p>This Certificate of Approval is valid for the above term provided there is no change in the protocol.</p> <p>Modifications To make any changes to the approved research procedures in your study, please submit a "Request for Modification" form. You must receive ethics approval before proceeding with your modified protocol.</p> <p>Renewals Your ethics approval must be current for the period during which you are recruiting participants or collecting data. To renew your protocol, please submit a "Request for Renewal" form before the expiry date on your certificate. You will be sent an emailed reminder prompting you to renew your protocol about six weeks before your expiry date.</p> <p>Project Closures When you have completed all data collection activities and will have no further contact with participants, please notify the Human Research Ethics Board by submitting a "Notice of Project Completion" form.</p>	
Certification	
<p>This certifies that the UVic Human Research Ethics Board has examined this research protocol and concluded that, in all respects, the proposed research meets the appropriate standards of ethics as outlined by the University of Victoria Research Regulations Involving Human Participants.</p>	
<p>_____ Dr. Rachael Scarth Associate Vice-President Research Operations</p>	

Certificate Issued On: 28-Jul-16

16-275 Minshall, Simon

Appendix 2: Invitation to Participate

Good Day,

You are invited to participate in a study entitled “Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records” which is being conducted by Simon Minshall, a graduate student at the School of Health Information Science at the University of Victoria.

The study has been approved and assigned Protocol Number 16-275 by Human Ethics Review Board at the University of Victoria.

The purpose of this research project is to develop an understanding of mobile devices including their ease-of-use and safety in health care. Research of this type contributes to knowledge of device use and patient safety in health care.

You are being asked to participate in this study because you are a medical doctor or a nurse practitioner and are well positioned to comment on the pragmatic aspects of these devices’ usefulness and their safety within the context of health care.

The research will be conducted by way of a clinical simulation. You’ll start with providing some demographic information. You will then use a health-records app on a smartphone to look up some hypothetical patient information and from that provide your

impression of the state of the patient's health. A brief interview about the experience follows. The sessions will be screen-recorded and videotaped.

If you are interested in this study or have further questions, you may contact the researcher via email to obtain additional information. Participation in this study is voluntary. Thank you for your consideration.



Simon Minshall

Appendix 3: Participant Consent Form

You are invited to participate in a study entitled Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records that is being conducted by Simon Minshall.

Simon Minshall is a graduate student in the department of Health Information Science at the University of Victoria and you may contact him if you have further questions by emailing [**email address**].

As a graduate student, I am required to conduct research as part of the requirements for a degree in Health Information Science. It is being conducted under the supervision of **Dr. Andre Kushniruk**. You may contact my supervisor at [**phone number**].

The study has been approved and assigned **Protocol Number 16-275** by Human Ethics Review Board at the University of Victoria.

Purpose and Objectives

The purpose of this research project is to examine the different sizes of mobile handheld devices such as tablets and smartphones to investigate the effect that the size of the device has on the usefulness and the problems encountered while using it to retrieve medical records.

Importance of this Research

Research of this type is important because people make mistakes. When physicians make mistakes and when they do so with patient information there is the potential that patients may be harmed. This research will attempt to measure errors and offer insight as to what kind of errors they make when using mobile devices of different screen sizes.

Participants Selection

You are being asked to participate in this study because you have professional experience with health records and are well positioned to comment on the pragmatic aspects of these devices' usefulness and their safety within the context of health care.

What is involved

The research will be conducted by way of a clinical simulation at a location convenient to you such as your clinic or offices. During the 60-minute session you will start with providing some demographic information. You will then use a health-records app on a smartphone to look up some hypothetical patient information and from that provide your impression of the state of the patient's health. A brief interview about the experience follows. The sessions will be screen-recorded and videotaped with your permission. Audio-tapes and written notes will be taken and a transcription will be made. Video and audio recordings will not be shared publicly.

Inconvenience

Participation in this study may cause some inconvenience to you, including giving up an hour of your schedule.

Risks

There are no known or anticipated risks to you by participating in this research.

Benefits

The potential benefits of your participation in this research include adding to the state of knowledge. Findings have the potential to affect procurement of mobile devices to clinical staff.

Voluntary Participation

Your participation in this research must be completely voluntary. If you do decide to participate, you may withdraw at any time without any consequences or any explanation.

If you do withdraw from the study your data will be used only if participant gives permission.

Anonymity

In terms of protecting your anonymity in this study you will be identified only by your participant ID and never by name.

Confidentiality

Your confidentiality and the confidentiality of the audiovisual data will be protected by strong encryption. The encryption key will itself be encrypted and will be kept separately and securely from the data. Paper documents created during the sessions will only have a numerical participant identifier.

Dissemination of Results

It is anticipated that the results of this study will be shared with others in the following ways: thesis, published article and presentations at scholarly meetings.

Disposal of Data

Data from this study will be disposed of by shredding, and the encrypted data will be electronically deleted from storage. This information will be securely stored for the life of the project plus one year.

Contacts

Individuals that may be contacted regarding this study include Simon Minshall and Dr. Andre Kushniruk

In addition, you may verify the ethical approval of this study, or raise any concerns you might have, by contacting the Human Research Ethics Office at the University of Victoria (250-472-4545 or ethics@uvic.ca).

Appendix 4: Candidate Form

Name: _____ Email: _____

Phone: _____ Job Title: _____

This is an invitation for you to participate in a study examining the relationship between form factor and usability in handheld mobile devices from a clinical health perspective. The study titled "*Evaluating the effect of display size on the Usability and safety of a mobile handheld application for accessing electronic medical records*" intends to provide answers to the following:

1. Is there a correlation between the size of the display and error rates observed?
2. What are user preferences for device type?
3. What types of problems do users have with the smartphone and tablet interfaces to an electronic medical record?

The study will take approximately 90 minutes of your time in a hands-on session using the latest technology followed by a brief interview. You will also be asked to complete a short demographic questionnaire and an interview by telephone or in person depending on location and availability.

If you consent to participate in this study, your identity, as well as your organization's identity, will be kept strictly confidential. Your name or any other identifying information will not appear in any presentations or publications from this research. Any identifying information collected during this interview will be removed. Published results will contain only anonymous data.

Have you ever used a
smartphone or a tablet computer?
If so, for how long:

Have you ever used or are you
currently using an
Electronic health Record system
(EHR)? If so, for how long:

Do you know another person who
would be willing to participate in
this study?

Appendix 5: Participant Session Scripts

Introduction and training script

“I’m going to ask you to review some patient data and provide me with an impression of the patient’s health.”

“The first thing we’ll do is introduce the app and walk through the tutorial pages. Use the [mobile device], open the [app] and tap on the help button along the bottom. Please read each paragraph aloud and swipe the images to navigate through the tutorial. When you’re finished, we’ll move on to using the [app]. Please remember to read the tutorial aloud and to think-aloud any comments or thoughts that may occur.”

[Before proceeding, allow the participant time to finish the tutorial]

Session script

[Reiterate the need for thinking aloud and demonstrate it if necessary based on their performance in training. Prompt the participant to locate the first case]

“Please use the app to locate Patient [patient name] and begin reviewing the patient’s data. Please continue to think-aloud during your investigation and be prepared to

summarise your impression of this patient's health when you are complete. What do you think is wrong with them?"

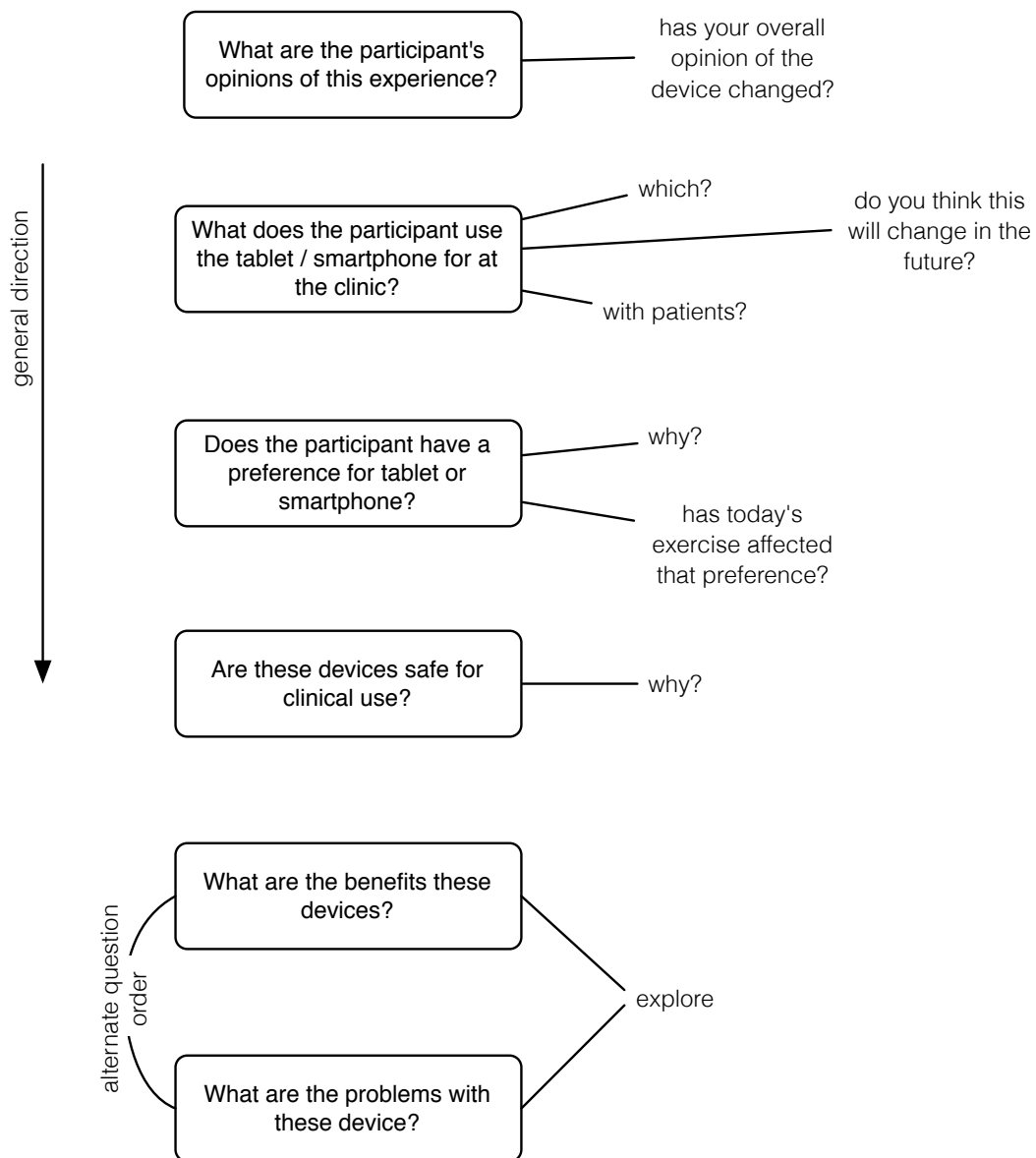
"The first task is to examine the watch-list on the home page. Using the 'Vitals' information please vocalise your impressions of what you read about the patient in case a. If you accidentally tap away from this area, please return to it and continue until you have reviewed all the patients in the list."

[If necessary, assist and prompt the participant to think aloud]

"From the home screen, please locate the patient and review their full charts. Please verbalise your impression and any finding and don't forget to think aloud while you are working."

[Offer assistance if needed and prompt to read-aloud]

Appendix 6: Semi-structured interview guide



Appendix 7: Technique for wireless screen recording of the iPhone® and iPad®

The specifics of setting up a computer for this task are constantly changing. In this guide, you will find an overview of what general steps to take and what to expect when things are working. Following the overview is a general setup guide that is intended to be read as a guide for an experienced researcher to configure a session. The list of specific steps need to get it all working is brittle, i.e., most likely to change as operating systems and other software continue to be updated.

For a research project using this method with many researchers, this guide should be used to create a walkthrough to be followed by each researcher. For the duration of the study, software versions should remain static and automatic updates prevented where possible.

Equipment Needed:

- A laptop computer equipped with:
 - A wireless network
 - A built-in microphone
 - Airplay client software (<http://www.airsquirrels.com/reflector/>)
 - Sufficient available storage for screen recording (20GB / participant).
- A WiFi router
- An iPhone® 4 and an iPad® 3 or newer (older models cannot screencast), with:
 - the app installed
 - a full battery
- An iPad® 3 or newer (older models cannot screencast) with the app installed

Setting:

The researcher sits across from the participant so that the participant does not get distracted by the movement of their action repeated on the laptop's display. They should be allowed to sit naturally and have enough room to deal with paperwork as the session progresses. With this setup the participant is required to remain in place during the screen recording. With some modification, i.e., a chest-mounted camera, this method could be adapted to allow perambulation during the session.

Overview:

The goal is to show both mobile devices' screens on the laptop screen. One of the devices is to be used by the participant to work with the app being tested. The other device is used as a camera to record the movements of the participants hands on the other device. This is useful because the mobile UI provides no mechanism for the researcher to see what part of the screen it being touched. Once both screencasts are shown on the laptop's display, use screen-recording software to record the whole display and set it to also capture audio from the laptop's microphone.

General Setup:

For the remote screen-casting to work, the two devices and the laptop must be connected to each other on the same wireless network. *The wireless network must allow peers to communicate to each other.* Most public WiFi installations do not allow peer-to-peer communications. Because of this it is preferable to create your own private wireless

LAN (WLAN) using a common wifi router⁵. Then, once the router is configured, power it up and the WLAN should be ready to use and you can set each device and the laptop to connect to its network. If the iPhone®, iPad®, and router are dedicated to this task, then this setup needs only happen once. Before a session begins, power the router on, check each device is connected to it.

The iPhone® and iPad® operating system allows for the screen to be broadcast, or, screen-casted, to a remote device with a technology named “Airplay”. This requires a suitable destination and as of this writing the laptop does not have a suitable Airplay client. To remedy this, you will need to install one. For this study, the author used *Reflector* (<http://www.airsquirrels.com/reflector/>). Once it’s installed, and configured according to the manufacturer’s guides, the laptop will appear as a destination. Set the iPhone® and iPad® to send its whole screen to the Airplay client and they should appear on the laptop’s display. Give one device to the participant and set the other one to record from its camera pointing it at the participant’s screen – it will make its own recording, which you can choose to keep – and also be shown and recorded on the laptop’s screen.

Initiate the screen recording and start the session after testing that you can see each device on the laptop’s display. At the end of the session stop the recording to avoid filling up your storage and its concomitant system problems, disable the screen casting, and stop the camera on the device that was recording the participant’s hands. Disconnect the router and the recording session is completely shut down.

⁵ It is also possible to have the laptop create its own WLAN and let the devices connect with it rather than to use an external router. The networking involved is the same but in the author’s experience the setup is cumbersome and needs to be repeated for each new session.

The participant's use of the iPhone® and iPad® apps

An analysis of the training data in the app shows that Mr. Hodges and Mr. Mills are similar cases. Both are male, in their 50s, admitted for hypertension, and have imaging, labs and other documents. In this Appendix Mr. Hodges information will be browsed using the iPhone® and Mr. Mills will be browsed with the iPad® as illustrated in the following figures and captions:

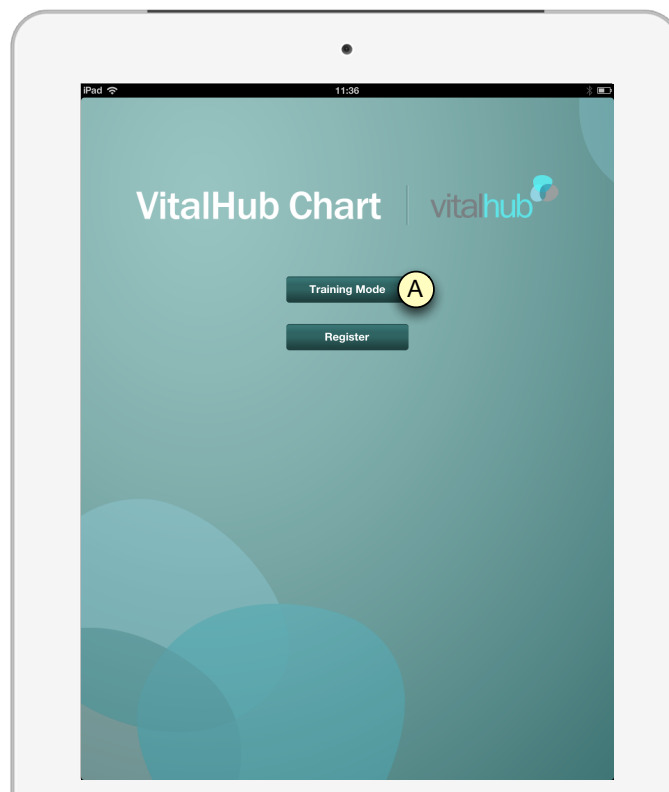


Figure 29: The Welcome Screen on the iPad®

When the app loads, the user is presented with this screen. Tapping the *Training Mode* button (A) navigates to the main user interface shown in Figure 30. In tablets and button-less smartphones, the user does not “click” as one would click with a mouse, rather, the user uses “taps”, “swipes” and “drags” on the screen in order to affect an interaction with the device. Moving from one distinct screen to another is termed a segue.

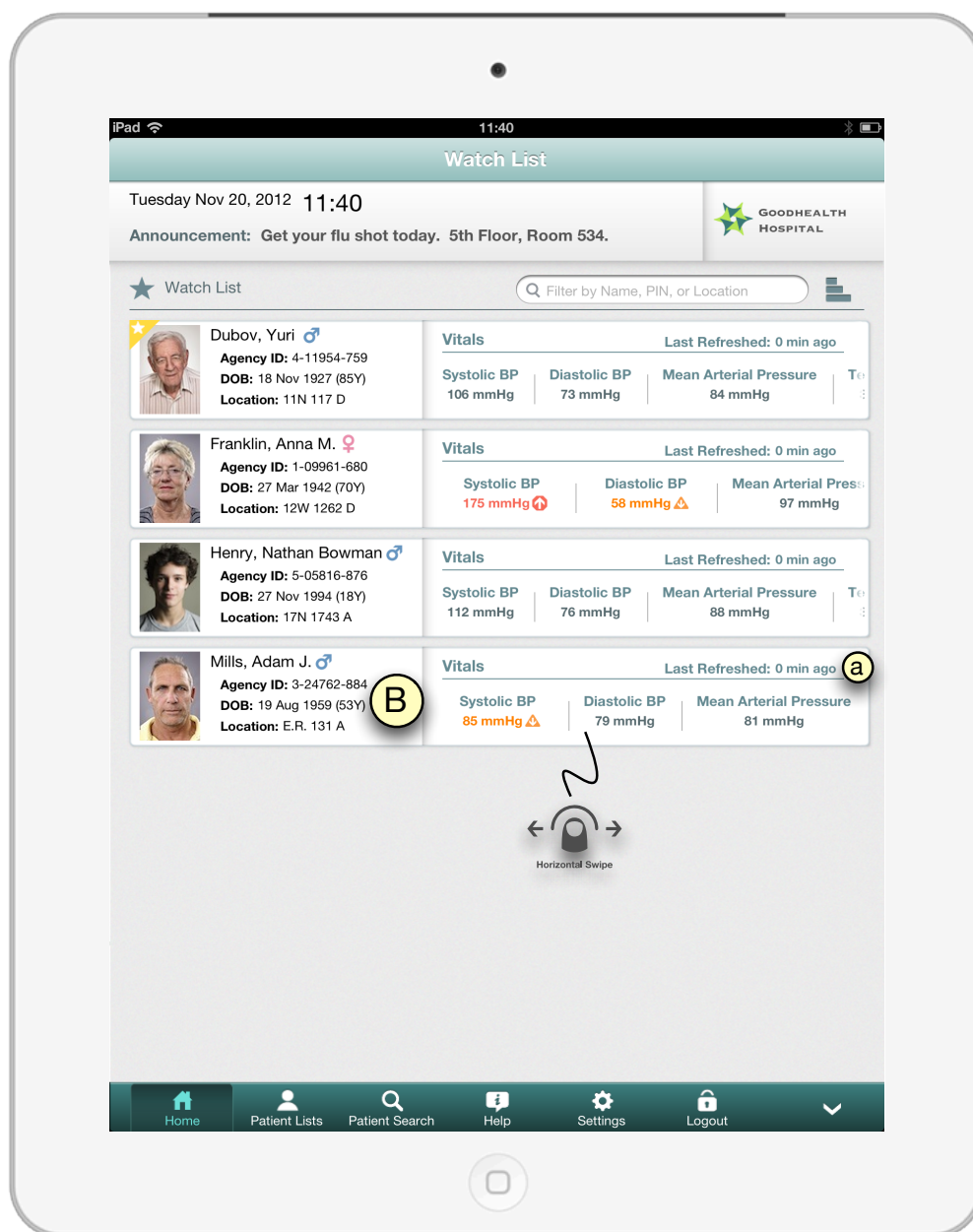


Figure 30: The VitalChart main iPad® user interface

The main interface shows the *Watch List* when it first loads. Tapping on *Patient Lists* pops up a dialog that allows the user to choose alternate lists, but in this case this is the list that the user wants. Mr Mills, our target patient, is shown on the bottom (B). His *Vitals*, and those of the other patients in this list are shown to the right of the demographic information. Swiping left-right on the *Vitals* scrolls to reveal all the detail the for patient. Tapping on the demographic area segues to a screen called the *Patient Dashboard* where the user can find much more summary data and from which navigation to the details is possible. The user taps the demographics area.

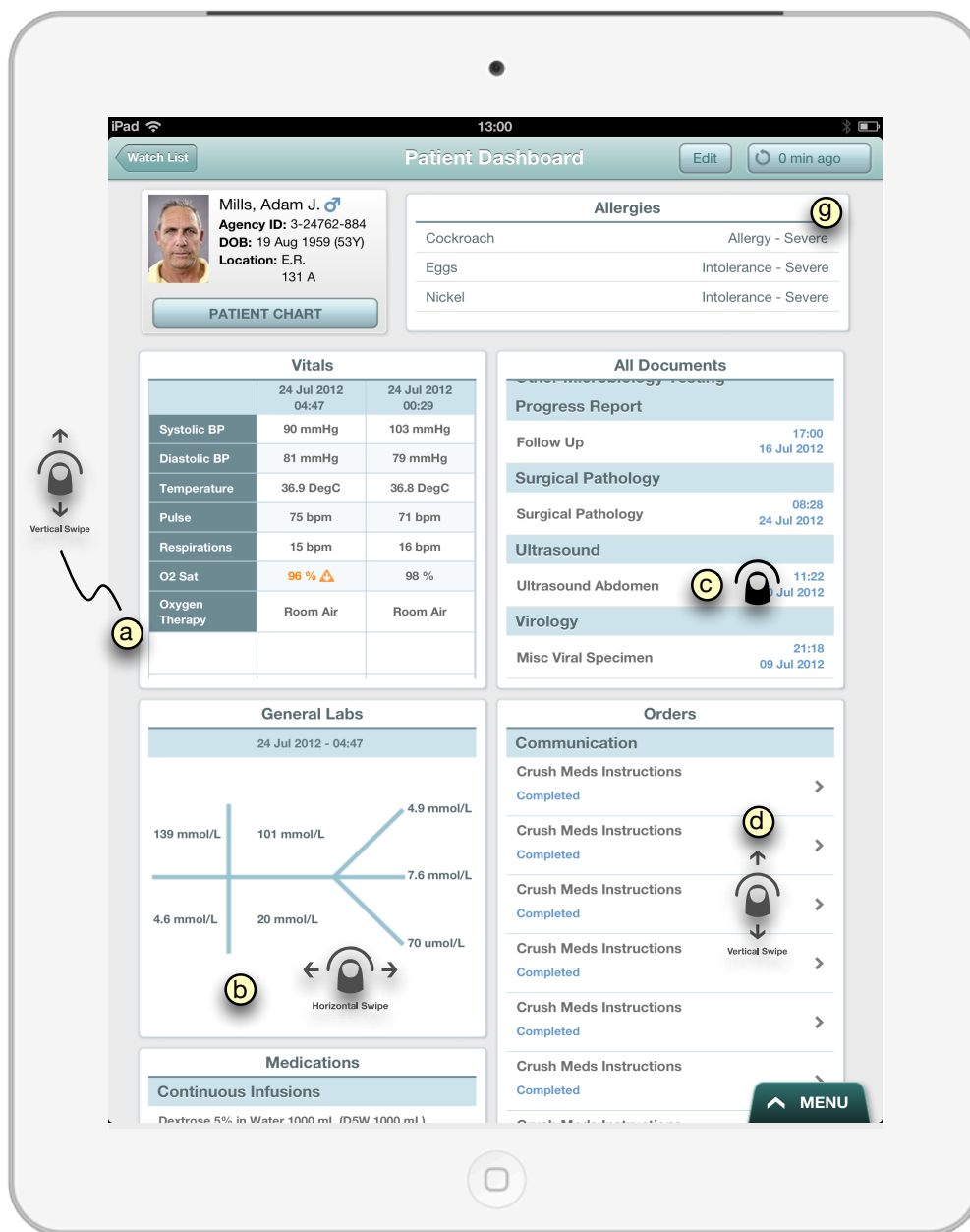


Figure 31: VitalChart's iPad® patient Dashboard

The Patient Dashboard displays all the summary data. The entire Dashboard, organised into separate panels, is too large to fit on the screen. The user scrolls up-down to allow the user to see all the panels. Each panel contains one kind of information. In panels where there are multiple records in a timeline, General Labs (b) and Vitals for example, swiping left-right will scroll through their records. In panels where the contents are too numerous, Documents and Orders (d) for example, swiping up-down will scroll through the contents so that the user has access to all the data. Swipes on panels take precedence over Patient Dashboard swipes making navigation somewhat awkward. The user is expected to tap on some panels items to view the detailed records, for example, tapping on (c) will segue to a view of the imaging report itself.



Figure 32: An example of a document on the VitalChart iPad® App

This is an example of an imaging report complete with the medical image, and report details displayed full-screen. The document is not editable but using a “spread” gesture by using two fingers on the screen and spreading them open while in contact with the screen, the document may be enlarged and repositioned for a clearer view. This feature may be used in either portrait or landscape orientation. The buttons at the top navigate back to the Dashboard and there the user may choose other details to view and do this until the patient review is deemed complete.

On the iPhone® the participant is guided to review Mr. Hodges Case. This patient is of similar age and was also admitted for hypertension. He is in a different ward so further navigation is required, not simply because the watchlist does not appear by default, but also because the screen is smaller and the arrangement of the panels and UI elements differs.

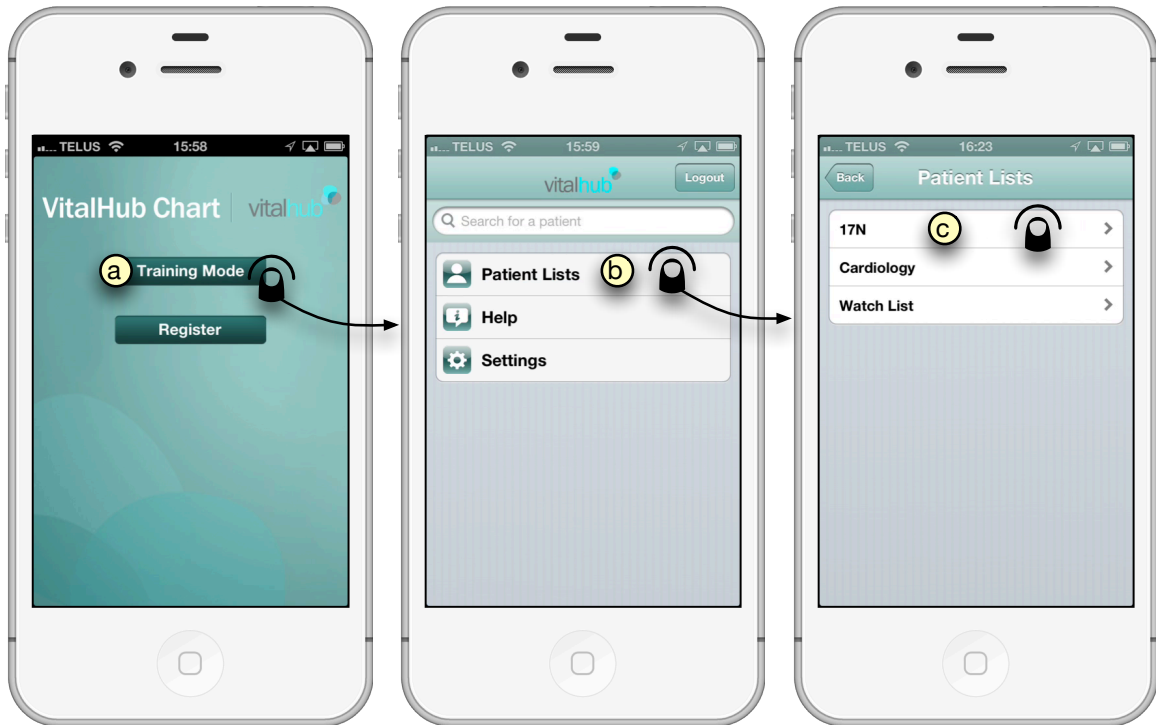


Figure 33: First three screens on VitalChart's iPhone® App

On the iPhone®, shown in Figure 33, the app has the same dataset (same patients, same health information) for its training mode. The model of panels and their contents is the same but because of the smaller screen size the navigation is different and most screens show only one panel or other UI element. As a result the user must increase the amount of interaction needed to get from the main page to the users details, in this case tapping on (a) segues the user to a home page where tapping on (b) segues to Patient Lists. Mr Hodges record may be found in the 17N list. The user may either search for it from the main page or browse for it in the lists. Tapping on (c) will segue to the list of 17N patients where our target appears at the top of the page.

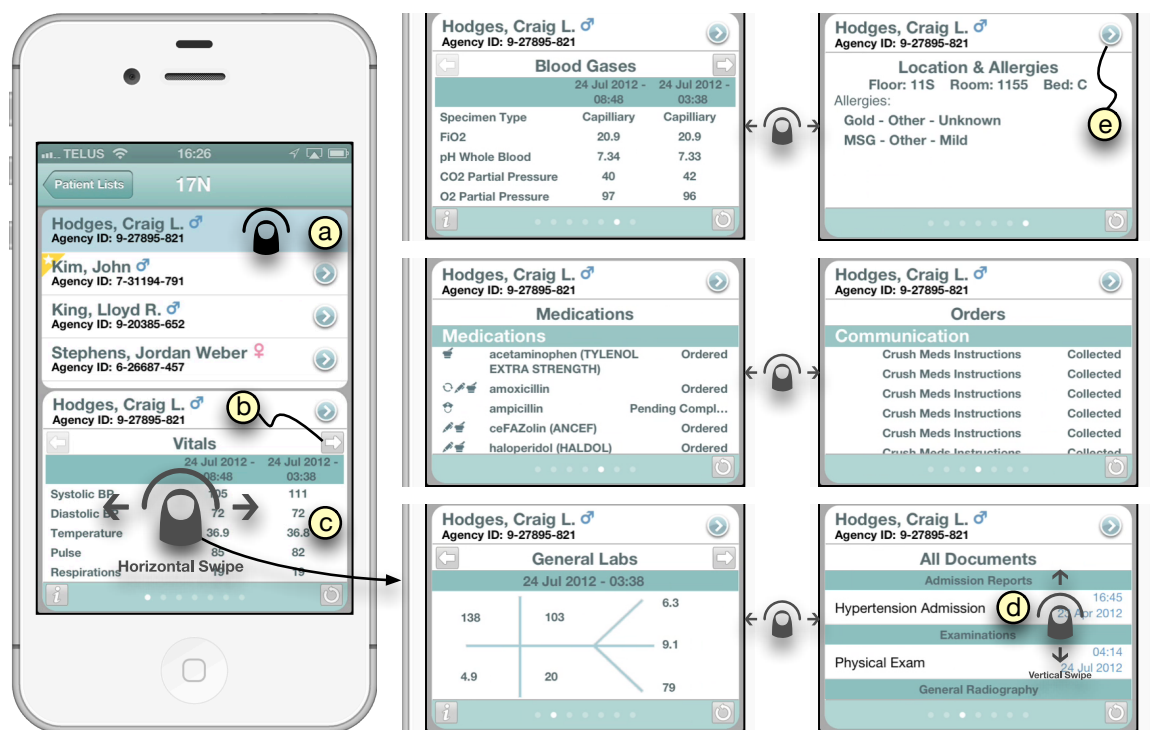


Figure 34: Patient data navigation on the VitalChart iPhone® App

Tapping on the Hodges line at the top of the screen (a) reveals Mr Hodges details at the bottom half of the screen, defaulting to the Vitals panel. Tapping the small arrow (b) scrolls through the panel's records. Swiping left-right in (c) scrolls through all the panels and the small set of navigation dots at the bottom of the screen gives the user an indication of where they are in the list of panels. Similarly to the iPad® interface, if a panel is too long to show all contents on one screen then the user may swipe up-down (d) to navigate all the data. Tapping on any of the disclosure icons, such as (e), segues to a full-screen view showing only the details for the currently selected patient.



Figure 35: Drilling down to details on the VitalChart iPhone® App

Tapping on the body (a) on some panels also segues to the full-screen view. This view allows the whole screen to be used to examine patient data. Horizontal and vertical swipes may be used where appropriate to navigate the data. This page shows record summary data. Choosing and tapping the disclosure icon on a particular piece of data (b) will segue to the full record details shown at (e).

When the review is complete, the user may return to the previous *17N* screen using the *Patient Lists* navigation button at the top of the screen (d) or remain in the full-screen view and jump to the other panels using the navigation bar (c) to jump to other panels and continue the review without returning to the *Patient Lists* page.

The next page, see Figure 36, is arrived at by tapping on the Documents icon on the navigation bar (c).



Figure 36: Navigating and viewing clinical documents on the VitalChart iPhone® App

Viewing the Documents panel in full screen view reveals the list of documents. Tapping the first X-Ray document (a) segues to the view of the actual document and its images and notes. Similar to the UI of the iPad® version, the document is not editable but using a “spread” gesture with two fingers on the screen and spreading them open while remaining in contact with the screen, the document may be enlarged and repositioned for a clearer view. The buttons at the top navigate back to the Dashboard and there the user may choose other details to view and do this until the patient review is deemed complete.

Appendix 8: Session Roster

The roster based on the final block for groups of four participants would look like this:

<i>Participant</i>	<i>Device</i>	<i>Case</i>
One	<input type="radio"/> iPad*	<input type="radio"/> <i>a</i>
	<input type="radio"/> iPhone*	<input type="radio"/> <i>b</i>
Two	<input type="radio"/> iPad*	<input type="radio"/> <i>b</i>
	<input type="radio"/> iPhone*	<input type="radio"/> <i>a</i>
Three	<input type="radio"/> iPhone*	<input type="radio"/> <i>a</i>
	<input type="radio"/> iPad*	<input type="radio"/> <i>b</i>
Four	<input type="radio"/> iPhone*	<input type="radio"/> <i>b</i>
	<input type="radio"/> iPad*	<input type="radio"/> <i>a</i>

(repeat)

Appendix 9: Demographics Questionnaire

Researcher: _____ Participant ID: _____
Location: _____ Date and time: _____

Instructions

Please answer each question as accurately as you are able.

Part I: Demographics

- 1) **What is your age?**
 - a. Under 24
 - b. 25 to 44
 - c. 45 to 64
 - d. 65 to 84
 - e. Over 85

- 2) **How long have you been a doctor or nurse practitioner?**
 - a. Started recently
 - b. Less than five years
 - c. Less than ten years
 - d. More than ten years

- 3) **How long have you been using a computer?**
 - a. Started recently
 - b. Less than five years
 - c. Less than ten years
 - d. More than ten years

- 4) **How long have you been using a smartphone or tablet?**
- a. Started recently
 - b. Less than five years
 - c. Less than ten years
 - d. More than ten years
-
- 5) **How long have you been using medical apps on a smartphone or tablet ?**
- a. Started recently
 - b. Less than five years
 - c. Less than ten years
 - d. More than ten years

Appendix 10: Post-case interview and questionnaire

Researcher: _____ Participant ID: _____
 Location: _____ Date: _____

Part 2a: Interview

5. Did you find this app easy to use for this case? Why or why not?
6. Did you find the app easy to navigate for this case? Why or why not?
7. Did you find the information easy to read for this case? Why or why not?
8. Did you find have any problems using the app? Please discuss?

Part 2b: Questionnaire

For each of the questions below, circle the number that best characterizes how you feel about the statement:

1. It is easy to use the app.

Strongly Disagree	1	2	3	4	5	6	7	8	9	Strongly Agree
-------------------	---	---	---	---	---	---	---	---	---	----------------

2. It is easy to navigate the app.

Strongly Disagree	1	2	3	4	5	6	7	8	9	Strongly Agree
-------------------	---	---	---	---	---	---	---	---	---	----------------

3. It is easy to read the information presented.

Strongly Disagree	1	2	3	4	5	6	7	8	9	Strongly Agree
-------------------	---	---	---	---	---	---	---	---	---	----------------

4. It is easy to use the device.

Strongly Disagree	1	2	3	4	5	6	7	8	9	Strongly Agree
-------------------	---	---	---	---	---	---	---	---	---	----------------

5. It is easy to understand the information presented.

Strongly Disagree	1	2	3	4	5	6	7	8	9	Strongly Agree
-------------------	---	---	---	---	---	---	---	---	---	----------------

Appendix 11: Post-case interview and questionnaire

Researcher: _____ Participant ID: _____
 Location: _____ Date: _____

Part 3a: Interview

1. Did you find this app easy to use for this case? Why or why not?
2. Did you find the app easy to navigate for this case? Why or why not?
3. Did you find the information easy to read for this case? Why or why not?
4. Did you find have any problems using the app? Please discuss?

Part 3b: Questionnaire

For each of the questions below, circle the number that best characterizes how you feel about the statement:

- | | | |
|----|--|--|
| 1. | It is easy to use the app. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |
| 2. | It is easy to navigate the app. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |
| 3. | It is easy to read the information presented. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |
| 4. | It is easy to use the device. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |
| 5. | It is easy to understand the information presented. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |

Appendix 12: Final interview and questionnaire

Researcher: _____ Participant ID: _____
 Location: _____ Date: _____

Instructions

For each of the questions below, circle the number that best characterizes how you feel about the statement:

Part 4a: Interview

1. Would you say that the iPad® or iPhone® are difficult to use in clinical practice?
Why or why not?
2. What are your thoughts about patient safety while using these devices in clinical practice? Please discuss.
3. Are you confident that you would be able to retrieve and interpret patient data by using these devices? Please discuss.
4. Are the screen sizes appropriate to use in clinical practice? Why or why not?

Part 4b: Questionnaire

- | | | |
|----|--|--|
| 1. | The iPad® is easy to use in clinical practice. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |
| | | |
| 2. | I would use the iPhone® with this app in clinical practice. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |
| | | |
| 3. | I would use the iPad® with this app in clinical practice. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |
| | | |
| 4. | The iPad® with this app is safe in clinical practice. | |
| | Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree | |

5. The iPhone® with this app is safe in clinical practice.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
6. There are advantages with iPhone® when using this app.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
7. There are advantages with iPad® when using this app.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
8. The iPhone® easy to use in clinical practice.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
9. It is easy to use this app with the iPhone®.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
10. I am confident I can retrieve the information I need using the iPhone®
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
11. The iPhone® is too small for clinical use.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
12. It is easy to use this app with the iPad®.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
13. Using the iPad® when working with patients is unsafe.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
14. Using the iPhone® when working with patients is unsafe.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
15. I would avoid using the iPad® in clinical practice.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

16. I would avoid using the iPhone® in clinical practice.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
17. The iPad® is too large for clinical use.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
18. I am confident I can retrieve the information I need using the iPad®.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
19. This app is not safe for use in a clinical setting.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree
-
20. This app is safe for use in clinical team meetings.
-
- Strongly Disagree 1 2 3 4 5 6 7 8 9 Strongly Agree

Part 4c: One Final Question

Which device did you like better? (please circle your answer)

iPhone®

iPad®

Appendix 13: Sample Coded Transcript

Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records

Coded Transcript

Participant Two: P2_screen_capture.mp4

Notes

See guidelines in Coded_Transcript_Notes.docx

Transcription

Start of the device-A (ipad, Swofford)

03:23

Researcher (R)

You need to go and find the patient named Norman Swofford. Think aloud

Participant 2 (P2)

Looking at this patient's Lab work. He has sepsis. Maybe that's his diagnosis. Maybe that's why he's in the hospital. It shows his blood gases and documents including an Admission Report dated 2015 and 2016.

03:28

P2

What I know about this guy so far is that he's febrile and his O2 sat is a little high so he's probably on supplemental oxygen and is getting a little too much. There is something wrong with his urine. oh there's protein in his urine. Though there's no indication in the urinalysis piece because there's no date when they were taken. So that's a little weird. His blood gases are off a little bit but not terrible. I can tell you what medications he's on.

2001 → USABILITY -- CONSISTENCY -- "NO DATES FOR DATA" -- SWOFFORD -- IPAD -- LABS -- FIRST_SESSION

→

Coded Transcript: Participant Two
Media File: P2_screen_capture.mp4

Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records

4:41

R

Which dates are you talking about?

P2

Yeah the other sections have dates on them but the Uranalysis part doesn't that's all.

R

You may be able to tab on a value to learn more

P2

OK, let's see. Umm he's got a history of Diabetes. It looks like this is a guy with Diabetes who's been admitted to hospital with sepsis which could potentially be coming from his urinary system. White blood cells in his urine. Glucose in his urine. Nitrate is positive so just at a glance I'd say he's got urosepsis which is a pretty common way to get sick with an infection.

5:45

I don't know how you back-up because would be helpful to see his ... ok so I can go back in

2002 → SAFETY_TIE -- SLIP -- "CANNOT NAV BACK" -- SWOFFORD -- IPAD -- LABS -- FIRST_SESSION

P2

Just to read his admission history to find out why he's at the hospital.

Ok so now we're into his physical exam which I would not normally start with.

Oh it's not obvious how to get back out of this there should be a way ...

Now the admission report looks like an outpatient visit it's not an admission report for acute care. But it does look

Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records

like from the other information that he was admitted for acute care. They call it an admission history but I'd say it was more like a clinical note. So there isn't one. OK what next?

- 2003 → USEFULNESS -- CORRECTNESS -- "NOT AN ADMISSION REPORT" -- SWOFFORD -- IPAD -- DOCUMENTS -- FIRST_SESSION
- 2004 → USEFULNESS -- APPLICABILITY -- "WHY'S PT IN HOSPITAL" -- SWOFFORD -- IPAD -- DOCUMENTS -- FIRST_SESSION
- 2005 → SAFETY_TIE -- SLIP -- MISNAVIGATION -- SWOFFORD -- IPAD -- DOCUMENTS -- FIRST_SESSION

R

Lets look at the row of icons along the bottom and see if we've covered everything.

7:31

P2

Yeah, ok let's see if there's any diagnostic imaging.

That's pretty good. I'm no radiologist but there doesn't seem to be any pneumonia in there. That's actually a pretty nice feature there. Our EMR doesn't easily show us imaging.

- 2006 → USEFULNESS -- IMPACT -- COMMENT -- POSITIVE -- SWOFFORD -- IPAD -- DOCUMENTS -- FIRST_SESSION

R

I noticed there that the image got brighter there. Is that happening on your screen?

P2

Yeah it did that as I touched the screen. I guess with reading xrays that contrast is everything.

Nothing there that's interesting to me

Just seeing if there is anything else ... yeah so ... perfusion test I'm going to

Coded Transcript: Participant Two
Media File: P2_screen_capture.mp4

Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records

have a look at that. Uh so if he's in hospital now, it was it was ordered on December 20th so if this person's in hospital now it's unrelated.

¶

▪ **R**
 BEGIN ipad POST-USE INTERVIEW
 09:27

▪ **R**
 Did you find this app easy to use for this case?

▪ **P2**
 Yes I would say I did. Well I find there's a good balance of information on the screen

▪ 2007 → USABILITY -- LAYOUT -- COMMENT -- POSITIVE -- SWOFFORD -- IPAD -- INTERVIEW_A -- OVERALL
 ▪ 2008 → USABILITY -- EASE_OF_USE -- COMMENT -- POSITIVE -- SWOFFORD -- IPAD -- INTERVIEW_A -- OVERALL

There's something about the layout that's easy to read. The font is easy to read and the balance between light and dark helps make it clear to see who the patient is.

▪ 2009 → USABILITY -- FONT -- COMMENT -- POSITIVE -- SWOFFORD -- IPAD -- INTERVIEW_A -- OVERALL
 ▪ 2010 → USABILITY -- GRAPHICS -- COMMENT -- POSITIVE -- SWOFFORD -- IPAD -- INTERVIEW_A -- OVERALL

There was a place in the Ureanalysis section that makes it unclear, like there is a column or row missing like the dates so you cannot tell when the data was taken.

▪ 2011 → USABILITY -- CONSISTENCY -- "NO DATES FOR DATA" -- SWOFFORD -- IPAD -- INTERVIEW_A -- LABS

10:43

Coded Transcript: Participant Two
 Media File: P2_screen_capture.mp4

Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records

R

Did you find the app easy to navigate for this case?

P2

I'd say pretty easy... yeah like maybe three quarters of the way

2012 → USABILITY → NAVIGATION → COMMENT → POSITIVE → SWOFFORD → IPAD → INTERVIEW_A → OVERALL

R

Was the app easy to read? Legibility?

10:57

P2

Yeah it was easy to read.

2013 → USABILITY → FONT → COMMENT → POSITIVE → SWOFFORD → IPAD → INTERVIEW_A → OVERALL

R

Did you have any problems using the app?

P2 →

It's not super clear to me like the patient's location. Are they in hospital? Like I still don't really know what's really going on with the patient. I noticed that he has vitals taken at 8:30 in the morning on the second of January and at 3:55 am on the 22 December so those are really unusual things so I don't have a full handle on why they're there and whether they're an inpatient or an outpatient. If they're an inpatient then the medication doesn't make any sense.

2014 → USEFULNESS → RELEVANCE → COMMENT → NEGATIVE → SWOFFORD → IPAD → INTERVIEW_A → VITALS

Page Break

Evaluating the effect of display size on the usability and the perceptions of safety of a mobile handheld application for accessing electronic medical records

Start of the second device (iPhone)

- 13:00
- R**

Look for a Jamila McKim. We'll use this iPhone and see what we can learn on the iPhone about the state of the patient's health.
 - P2**

Looks like they've been in Hospital for about 24 hours - or at least in contact with whatever system this is.

13:42

oops. It's hadr to back up. I meant to mention that on the last system. So I don't really know how to get to the next section.
 - 2015 → SAFETY_TIE -- SLIP -- 'EXIT TO LIST' -- MCKIM -- SECOND_SESSION -- IPHONE -- VITALS**

 - R**

Yes, I see. The word Dashboard there is tappable and will pop up a list when you tap it.
 - P02 →**

Allright, so Documents is the one I'm gonna go for. Admission Reports is always a go-to place.

14:24

Oh, that's super-hard to read! [reads report details]
 - 2016 → USABILITY -- FONT -- 'HARD TO READ' -- MCKIM -- SECOND_SESSION -- IPHONE -- DOCUMENTS**

 - P2**

...ok so that's terrific but doesn't really jive with the clinical thing.
 - 2017 → USEFULNESS -- APPLICABILITY -- MCKIM -- SECOND_SESSION -- IPHONE -- DOCUMENTS**

Coded Transcript: Participant Two
Media File: P2_screen_capture.mp4