

Usability Methods for Ensuring Health Information Technology Safety: Evidence-Based Approaches

Contribution of the IMIA Working Group Health Informatics for Patient Safety

E. Borycki¹, A. Kushniruk^{1,2}, C. Nohr², H. Takeda³, S. Kuwata⁴, C. Carvalho¹, M. Bainbridge¹, J. Kannry⁵

¹ School of Health Information Science, University of Victoria, Canada

² Department of Development and Planning, Aalborg University, Aalborg, Denmark

³ Jikei Institute, Osaka, Japan

⁴ National Cerebral and Cardiovascular Center Hospital, Osaka, Japan

⁵ Mount Sinai Medical Center, New York, New York

Summary

Objectives: Issues related to lack of system usability and potential safety hazards continue to be reported in the health information technology (HIT) literature. Usability engineering methods are increasingly used to ensure improved system usability and they are also beginning to be applied more widely for ensuring the safety of HIT applications. These methods are being used in the design and implementation of many HIT systems. In this paper we describe evidence-based approaches to applying usability engineering methods.

Methods: A multi-phased approach to ensuring system usability and safety in healthcare is described. Usability inspection methods are first described including the development of evidence-based safety heuristics for HIT. Laboratory-based usability testing is then conducted under artificial conditions to test if a system has any base level usability problems that need to be corrected. Usability problems that are detected are corrected and then a new phase is initiated where the system is tested under more realistic conditions using clinical simulations. This phase may involve testing the system with simulated patients. Finally, an additional phase may be conducted, involving a naturalistic study of system use under real-world clinical conditions.

Results: The methods described have been employed in the analysis of the usability and safety of a wide range of HIT applications, including electronic health record systems, decision support systems and consumer health applications. It has been found that at least usability inspection and usability testing should be applied prior to the widespread release of HIT. However, wherever possible, additional layers of testing involving clinical simulations and a naturalistic evaluation will likely detect usability and safety issues that may not otherwise be detected prior to widespread system release.

Conclusion: The framework presented in the paper can be applied in order to develop more usable and safer HIT, based on multiple layers of evidence.

Keywords

Usability testing, clinical simulation, technology-induced error, health information technology, patient safety

Yearb Med Inform 2013;20-7

Introduction

Ensuring health information technology (HIT) usability and safety has become an important goal for system designers and implementers. However, issues related to a lack of system usability continue to be reported in HIT [5, 13, 14]. The impact of poor usability on the safety of HIT has come to the fore over the past several years. More specifically, HIT usability has become a safety and quality issue for healthcare policy makers, administrators and vendors. In this paper, we describe a range of methods that have evolved from the empirical literature arising from usability engineering that can be applied to developing safer HIT prior to systems implementation. In addition to this, our paper describes how a combination of standard usability techniques have been extended to make HIT safer in an effort (by researchers) to mitigate and reduce the risk of releasing potentially harmful IT to be used by health professionals in the process of patient care. In this paper we: (a) provide a brief overview of the definitions for usability and technology-induced error, (b) outline how usability engineering techniques have been extended to cover HIT safety, and (c) describe a framework for applying usability engineering methods for identifying potential technology-induced errors. We begin our paper by defining usability.

What is Usability?

Usability refers to a measure of ease of use of IT. Usability of a healthcare system can be considered in terms of the following dimensions: effectiveness (does the system

do what the end user wants?), efficiency (how proficient is the system in carrying out user tasks?), enjoyability (how pleasant or satisfying is the system to use?), learnability (how easy is it to learn how to use the system?) and safety (how safe is it to operate the system?) [2, 5, 15, 30]. Over the past several decades, the usability of HIT has come to be recognized as a critical and significant issue in the success or failure of HIT (as it has been shown to be directly related to how usable systems are) [11, 23]. Successful implementation and adoption of HIT depends on the good usability of a technology, and problems with usability have been implicated in the failure of end users to adopt a range of HIT applications [16]. In addition, research has shown that poor usability of critical systems in the healthcare domain is also statistically related to medical errors by health professionals (i.e. technology-induced errors) [5] – with an increasing number of reports of patient harm and even death [14, 17]. Technology-induced errors may arise from the use of HIT if not designed and adequately tested from a usability perspective [5, 18].

What are Technology-Induced Errors?

Technology-induced errors involving HIT arose as a significant and important issue in the HIT literature in 2004-2005 [5, 13]. Since then a number of research publications emerged describing examples of such errors leading to patient harm, disability and/or death [14, 17] from researchers in Australia [14, 17], Canada [1, 5, 19, 20], the United States [11, 21], Japan [22] and the EU [23, 24]. With this

international focus on technology-induced errors, an important and critical need to understand these new types of medical errors emerged [18]. Medical errors related to the use of information technology (i.e. technology-induced errors) come from “(a) the design and development of technology, (b) the implementation and customization of a technology, and (c) the interactions between the operation of a technology and the new work processes that arise from a technology’s use” [18, p.154].

Internationally, researchers have found that such errors have arisen from suboptimal requirements gathering, system design, implementation and maintenance of HIT systems [18]. Unlike many other more traditional software errors (found during software testing such as white or black-box testing), technology-induced errors are often unobservable until systems are used under real-world conditions and/or in real-world settings (e.g. hospitals, clinics) [5, 14, 17]. In an attempt to prevent these errors from occurring, researchers developed and extended many traditional usability engineering methods to be able to describe and diagnose the causes of technology-induced errors with the aim of preventing them from occurring [5].

Usability researchers began extending techniques from the usability engineering literature in an effort to identify and address errors prior to the release of HIT systems. Researchers found that these new types of errors had links to HIT interface design [5], poor clinical workflow [6] and poor integration with organizational policies, procedures and processes [6]. Specific methods developed from these extensions to traditional usability engineering techniques appeared to be promising in identifying technology-induced errors proactively – i.e. prior to systems implementation, e.g., usability inspection [26] and usability testing [5].

Since then, researchers have developed and fine-tuned these methods so that they can be used to detect technology-induced errors prior to systems release. Many of these methods are at the intersection of the usability and HIT safety literatures and have their theoretical and empirical origins in both the usability and industrial safety research. More specifically, these proactive methods for assessing usability and safety of HIT (prior to widespread system deployment) include: usability inspection methods [2],

clinical simulations [27, 28, 29], and usability testing [2, 5, 12], which have been shown to be able to identify potential HIT safety issues, where interface design and workflow are concerned [2, 12]. All of these methods represent extensions of traditional usability methods to this new research area at the intersection of technology-induced errors and HIT safety (See Figure 1). In the next section of this paper we outline these methods and how they have been used to improve HIT safety.

Usability Inspection Methods and HIT Safety

Usability inspection methods involve one or more usability analysts systematically stepping through (or “inspecting”) a user interface or system. One approach to usability inspection is known as heuristic evaluation and it involves comparing features and functions of the system being evaluated against a set of heuristics, or guidelines, of good interface design [30]. Using this approach, violations of heuristics are noted. Several researchers have used heuristic evaluation in HIT device design and evaluation [2, 31]. The other widely applied inspection method is known as cognitive walkthrough. In applying cognitive walkthrough, an analyst steps through the HIT user interface noting its goals, actions, system responses and potential user problems [32]. In healthcare, both heuristic evaluation and cognitive walkthrough have been applied to the testing of a wide variety of systems and devices.

Guidelines have been developed to support the design of safer HIT by vendors (e.g. the Common User Interface Guidelines). These guidelines have arisen from healthcare and information technology industry partnerships such as the one between the National Health Service in the United Kingdom and Microsoft®. The Common User Interface Guideline program has led to the development of a “portfolio of standards and guidance related to the design of user interfaces for healthcare computing systems” to “increase patient safety, the clinical take-up of HIT and reduce training costs”. Researchers and vendors have partnered with the NHS and Microsoft® in extending and implementing these guidelines [33]. Guidelines have also been used to im-

prove the safety of HIT systems in terms of designing effective user interfaces [34].

Another trend has also emerged in the area of improving systems safety. This trend is concerned with the development of “evidence-based” safety heuristics or guidelines that can be applied to evaluating HIT in terms of potential safety issues and problems (for example, evaluating a computerized physician order entry (CPOE) system or electronic patient record). Carvalho et al. [35] undertook a multi-stage process where they developed safety heuristics, beginning with a systematic review of the literature, followed by an expert panel review of the literature and development of evidence-based heuristics (see Figure 2). After developing an initial set of evidence-based safety heuristics, researchers tested them conducting a heuristic evaluation and cognitive walkthrough involving a real-world health information system (i.e. the Veteran Affairs Computerized Patient Record System). Researchers found the heuristics were informative, providing safety information about varying static elements of the information system such as “system displays of medication status” and “system menus that are scrollable and clearly marked” [35, p. 52]. Alternatively, researchers did identify several limitations associated with the use of safety heuristics. They noted heuristics did not take into account local (organizational or country) guidelines, clinical processes or contexts. Furthermore, many of the safety heuristics could not be fully applied. System functions and underlying workflow processes could not be fully tested using the developed safety heuristics as representative patient cases, users and organizational conditions need to be present to fully understand the effects of organizational and country context on system safety. Researchers suggested the use of clinical simulations in conjunction with usability testing techniques for the full use and application of the guidelines to the dynamic aspects of HIT [35].

Therefore, HIT safety researchers have advocated conducting another layer of testing in addition to applying evidence based safety heuristics using heuristic evaluation and cognitive walkthrough – the use of clinical simulation testing. The combination of heuristic evaluation, cognitive walkthrough and clinical simulation can provide chief information officers with preliminary data for making initial

judgments about the safety of HIT during a system's procurement process and design or during its customization to a health care environment (e.g. hospitals, clinics).

In summary guidelines have been developed to drive the design of safer HIT. Some guidelines have been developed from a review of the literature and evidence in the area of HIT safety. Other guidelines have been effectively used to identify some aspects of HIT interface design and emergent clinical workflows that may lead to potential safety issues [35].

Usability Testing Methods and HIT Safety

In contrast to usability inspections (which do not involve studying end users), usability testing typically involves observing representative users of systems (e.g. physicians or nurses) as they carry out representative tasks (e.g. using an electronic health record, mobile health application) [2]. This type of testing is often conducted in a usability laboratory [3], where it has proved to be able to detect basic usability problems and to detect some types of technology-induced errors [2]. However, as typical usability testing does not involve testing systems under highly realistic conditions (i.e. in the actual setting of system use), certain types of errors may remain in the system after usability testing is completed and these errors are not detected until the system is in actual use under real-world conditions in real-world settings, which is too "late", particularly when patient safety is a concern. To address this issue, methods termed clinical

simulations have been developed that can be used in conjunction with usability testing for evaluating system safety. Clinical simulations when used in conjunction with usability testing allow for HIT to be tested within the context of real-world healthcare situations in response to the demands of real-world settings – e.g. urgent, atypical situations that are complex and may be life threatening [6].

Clinical Simulations and HIT Safety

A clinical simulation may be considered to be an extension of usability testing as it involves observing representative users carrying out representative tasks when using the system under study (which defines usability testing).

However, it extends evaluation because testing is carried out under representative contexts of use (i.e. in real or realistic settings and contexts) [4]. Clinical simulations have been found to be especially useful in detecting problems that appear in the interactions between a healthcare system and complex health professional work activities [5]. Clinical simulations are therefore a useful approach for detecting technology-induced error and have been identified to provide useful empirical evidence about how a system may affect information seeking [36], cognition, decision-making [5] and workflow [6, 27, 28].

Researchers have found that different types of problems may be detected by applying differing methods such as standard usability testing and clinical simulations. By applying laboratory style usability testing, evaluators can create scripts that "force" the user to go through

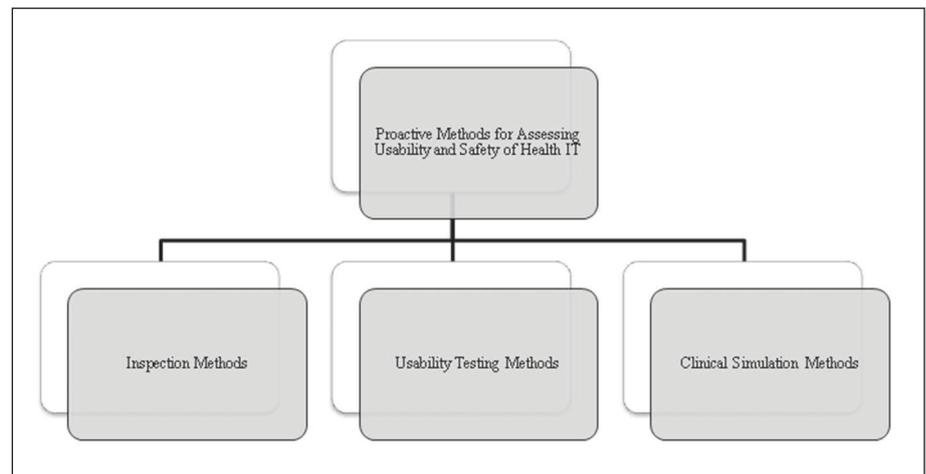


Fig. 1 Usability Methods and Health IT Safety

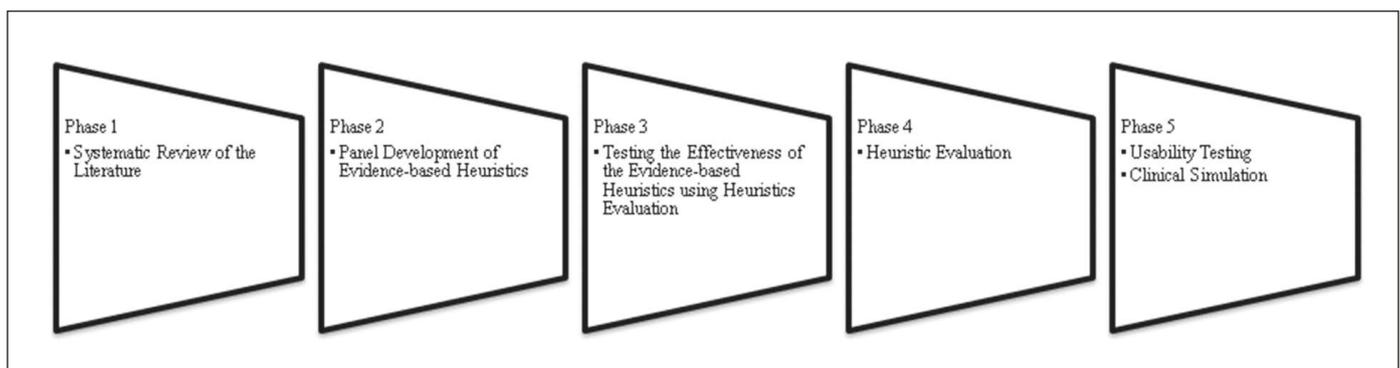


Fig. 2 Developing Safety Heuristics

certain components of a system or paths of interaction they would like to test. On the other hand, clinical simulations have the advantage of allowing users to interact with a system in a natural way that more closely reflects or resembles how they would use the tool in a real-world context. This may allow for detection of issues related to interface design and workflow as well as additional usability problems that may not be noted using traditional laboratory-style usability testing. Currently, there is an increasing desire for improving the level of evidence that HIT is safe and usability engineering methods are beginning to be applied more widely for this purpose. Such methods do not replace approaches such as focus groups, interviews, or post-implementation observation and error reporting. However, the sophistication of the application of usability engineering has reached a point whereby such methods can be employed proactively prior to system release to predict and prevent safety issues.

Indeed, according to a recent report by the Institute of Medicine [25], there is an increased focus on improving methods used in the design and implementation of HIT. The report suggests that usability engineering methods (in conjunction with clinical simulations) are increasingly being recognized as an evidence based approach that can provide data about technology-induced errors including frequency statistics.

A Framework for Applying Usability Engineering: Inspection, Testing and Clinical Simulation

In order to provide developers of HIT with evidence about the potential for their systems to cause technology-induced errors, a new trend is to integrate multiple approaches (i.e. usability engineering and simulation approaches described above) in new ways. In this part of the paper we describe a multi-phased approach to safety testing involving HIT that we have used as a framework to guide our system evaluations. The safety testing allows HIT to be developed and refined while at the same time preventing or mitigating the risk for technology-induced errors (see Figure 3).

We initially begin by having an analyst conduct a heuristic evaluation and cognitive walkthrough, which is followed by the correction of serious usability problems. This is followed by usability and clinical simulation testing of the system in laboratory and then “in-situ” (in real or highly realistic) settings, and finally may involve conducting naturalistic testing in-situ (with a limited number of users). The overall objective of the approach is to collect as much empirical evidence about the safety of HIT (prior to releasing it) as it is practically possible using evidence based approaches. This type of evidence is complementary to usability inspections and application of safety heuristics (as described in the previous section).

Phase 1 – Heuristic Evaluation and Cognitive Walkthrough

Data from an initial heuristic evaluation is used to identify potential interface features that may lead health professionals to make technology-induced errors. Here, analysts conduct heuristic evaluations and a cognitive walkthrough of the system or of a system prototype noting safety issues. Static elements of the interface design are identified that may lead to technology-induced errors [35] (recognizing that the dynamic aspects of the system will be evaluated for safety using clinical simulation testing) [20, 27-29, 36].

Phase 2 – Correction of Defects Identified by Heuristic Evaluation and Cognitive Walkthrough

Following this, data from Phase 1 is analyzed to identify interface design features and functions that may lead to technology-induced errors. This knowledge is used to refine or modify the interface to reduce the likelihood of technology-induced errors [35].

Phase 3 – Basic Usability Testing

This testing aims to identify and analyze basic or surface level usability problems. This involves having representative users of the system (e.g. physicians or nurses) carry out a set of scripted tasks under artificial laboratory conditions. In this type of testing the subjects are typically instructed to “think aloud” as they work through the scripts (e.g. for the entry of

data into an EHR or the use of a clinical guideline). Using this approach, computer screens are recorded, along with audio recordings of user’s thinking aloud while using the system or a system component (as described in detail in Kushniruk and Patel [2]). Audio data from think aloud sessions are transcribed, with the screen and video recordings used to annotate the transcriptions (using coding categories to identify specific types of usability problems, such as navigational problems, consistency issues etc. – see [2] for examples of coding approaches used to analyze HIT). Previous research has indicated that this type of testing, conducted with as few as 10-15 representative users can often identify the majority of serious basic surface level usability problems.

Phase 4 – Correction of Defects Identified by Usability Testing

Analysis of data from Phase 3 includes summarization of the number and frequency of basic level usability problems and potential technology-induced errors [5]. Based on a categorization of the frequency, impact, cost and urgency of fixing each identified usability problem, a set of recommended changes are made for improving or customizing the system. An important objective of this phase is to address the most significant problems identified in Phase 3 before moving to the next phase.

Phase 5 – Clinical Simulation

After base level usability problems have been corrected in Phase 4, our approach to testing involves moving to more realistic testing in real (or near real) settings and work contexts. This typically involves “in-situ” clinical simulations [6]. These simulations involve observing users using the system under study (e.g. an electronic health record) under simulated conditions that approximate real use. In our research, this has involved conducting observational studies in actual clinical settings (e.g. hospital rooms, operating rooms, intensive care units etc., after hours), with simulated patients (i.e. collaborators playing the role of patients), or in other cases using “digital patients” (e.g. involving use of interactive multimedia recordings of patients) [7]. This type of testing can be used

to assess how interactions with systems such as electronic health records are initiated by clinicians, and how the system affects clinician workflow. This testing can also be used to identify technology-induced errors that may only surface during live or “near-live” interactions (e.g. need for an emergency override that may arise only under certain work conditions). As in Phase 3, analysis consists of transcribing both user interactions with the technology under study, as well as audio recordings of verbal interactions with patients or other health professionals involved in the simulation.

Phase 6 – Correction of Defects Identified by Clinical Simulation

Data (i.e. audio and video recordings) from Phase 5 clinical simulations is analyzed to identify the following: (1) standard usability errors (which were not caught by “classic” usability testing in Phase 1), (2) inefficiencies and sub-optimal workflow patterns imposed by the system, and (3) technology-induced errors related to the complex interaction of the system user under realistic conditions of use. Significant problems are summarized and presented to the design or development team (depending on whether the system is being developed, or is being customized) for prioritization, with serious usability problems and potentially dangerous technology-induced errors rectified prior to the next and last phase. Although usability testing in Phase 3 may detect basic usability problems that may affect system safety, issues that may only be

detected by clinical simulations (e.g. effects of time-constrained or urgent healthcare work situations) may have a potential negative impact on system safety if not detected.

Phase 7 – Testing in Naturalistic Settings

When possible (given logistical and ethical constraints for doing so) we have recommended conducting limited naturalistic recording of use of HIT after Phase 6 changes have been implemented (and prior to widespread release of the system). Our experience has been that although a combination of usability testing followed by clinical simulation will lead to detection of errors that will not be “caught” by conventional usability-testing alone, it is impossible to predict through simulation all errors that may arise in a real setting. Therefore, in a number of studies we have additionally collected “live” data after clinical simulations are completed. This has typically involved using unobtrusive recording methods and technologies.

Phase 8 – Correction of Defects Identified by Naturalistic Testing

Prior to widespread system release, the recorded audio and video from Phase 7 are reviewed by the evaluators in order to identify potentially significant errors or need for customization before going fully “live” in a healthcare organization.

The overall evaluation framework we have developed using this phased approach to testing user aspects of HIT systems is de-

picted in Figure 3. Ideally, in order to improve the likelihood that major usability problems and technology-induced errors are detected (and ideally corrected) prior to widespread system release, all the phases shown in Figure 3 would be undertaken. The approach represents a “safety net” whereby usability errors missed from traditional testing may well be caught during the subsequent phases. Likewise technology-induced errors (e.g. inadvertent negative impact of a system on health professional workflow) may be detected during clinical simulation, but if they are not, then a final phase of limited naturalistic observation and testing may also “catch” problems prior to widespread release.

Usability Testing and Clinical Simulations Undertaken After Full System Deployment

It must be noted that ongoing monitoring for potential and actual safety issues needs to be undertaken after a HIT system has been released on a large scale within a healthcare organization. This work may involve reviewing near miss and actual incident reports (e.g. obtained through reporting mechanisms or verbal reports from staff). This may involve employing methods such as ethnographic techniques (including observations, interviews and focus groups with end users) after implementation and case study based approaches to understand where the errors

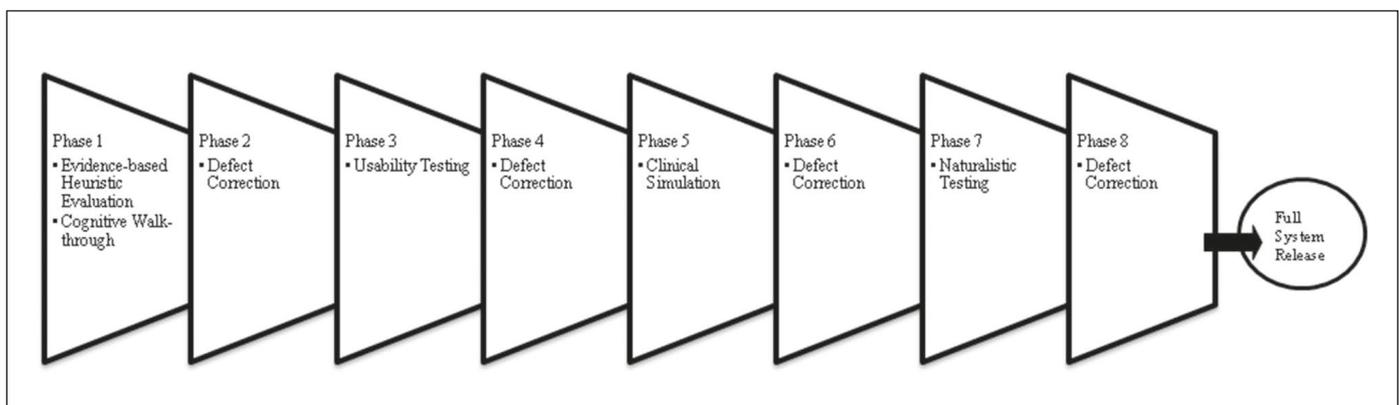


Fig. 3 Phases of System Testing Incorporating Usability Testing with Clinical Simulation and Naturalistic Testing

arose from. However, from our experience this work often needs to be complemented with targetted studies involving usability testing and clinical simulations, where errors can be observed in simulated settings and the activities and factors that lead to them can be better understood at a detailed level of observation. Such work is necessary to ensure usability and safety of systems over time and requires iterative cycles of testing and refinement of systems after initial full system deployment. There is also a need to conduct such work and to report and publish the findings in order to not only improve the safety of systems currently implemented, but also to develop a science of safe user interface design features and functions (as well as associated clinical workflows that lead to both usable and safe systems).

Time Constraints

All the phases described in the previous section identify differing types of safety issues and potential errors. In ideal world circumstances, they should be all undertaken to promote system safety. Researchers recognize that existing projects may not have allotted time for all activities. Due to time (e.g. project and implementation schedules timelines that have not taken into account the use of such phases) and economic constraints (e.g. if budgeting for testing was not taken into consideration in advance), we recommend that a key subset of the stages shown in figure 3 be conducted. At minimum we recommend undertaking activities from Phase 1 involving usability inspection (as heuristic evaluation and cognitive walk-through can be conducted rapidly and at low cost, often revealing important potential errors). Although usability inspection can detect some potential issues, dynamic elements of the interface design are more readily assessed applying usability testing as in Phase 3. We therefore recommend at a minimum that both usability inspection and usability testing methods be employed prior to releasing systems for real-world use. Regarding Phase 5 clinical simulations, where testing is conducted under more realistic conditions than usability testing, it is highly recommended that whenever possible this type

of testing is also conducted prior to releasing systems for real-world use (see the next section regarding economic considerations in doing this). Finally, it may not be possible to conduct studies recording live data in the naturalistic setting as in Phase 7. However, in life critical systems such testing should be considered. It is our hope that future projects will integrate more of these phases as knowledge of the importance of usability and safety grows.

Economic Considerations

In the past, testing of systems in clinical settings or simulation laboratories was considered difficult and expensive to do. However, researchers have developed portable low cost recording methods [see 4,7,12, 27] that can be taken into real-world settings and can be used to conduct both usability testing and simulation testing at a low cost (without having the cost of creating a simulated environment if these studies are actually conducted in examination rooms, hospital rooms and even operating rooms during off-hours). The use of existing rooms with equipment that is already being used by the organization for testing increases the ecological validity of studies. As well, the use of low-cost video recording equipment and free screen recording software can also significantly reduce costs [12].

Application and Experiences to Date

The approach and overall framework we have developed (and described in the previous section) was built on work over the last two decades by the authors in the design and evaluation of a wide range of HIT applications and systems. We have applied the framework to various degrees in testing and refining a wide range of HIT applications [2, 5, 6, 12, 27, 28, 36]. Historically, our initial application of traditional laboratory-style usability testing (in conjunction with usability inspection) has proven to be useful in a wide range of projects we have conducted in evaluating electronic health record systems, decision support systems and patient clinical information systems. However, despite the advantages of the approach as de-

scribed above over time, many of our projects have required the additional consideration of complex work environments in which systems we are testing are to be deployed.

In our first project combining usability testing with clinical simulation, the authors were requested to conduct user testing of a medication administration system being deployed in a hospital in Japan [6]. The aim of the evaluation was to ensure both the usability of the system as well as identifying any major technology-induced errors that the system might cause users (i.e. physicians and nurses) to make. Although an initial phase of usability inspection followed by traditional laboratory-style testing was found to be useful, it was clear that such testing would not provide a fully accurate picture of what would happen once the system was released for clinical use. Hence a further layer or “phase” of testing was initiated where users of the system were observed interacting with the system during clinical simulations. These clinical simulations were conducted in-situ in a typical hospital room where the system was to be deployed (the room was booked “off hours” for the study). The advantages of using this setting was the increased realism of the situation as well as the integration of the system with other interfacing technologies, such as bar-coding and scanning technologies (which will be interfaced with the system during real use). Subjects were asked to carry out a simulated set of medication administration tasks, interacting with a “dummy” patient and all interfacing technologies. Simulation cases were constructed, varying from simple drug administration, to more complex administration tasks (with a large number of medications, conflicting medications as well as simulated interruptions). This phase of testing revealed a set of potential usability problems and technology-induced errors not identified during the initial phase of standard usability testing. This included the need to customize the system to include an emergency override under certain conditions where the rigid and lock-step workflow of the medication administration system would pose a safety hazard. Other problematic aspects of the system were identified and corrected, including the identification of a problem where the system would lock other users out from a patient record if the user was interrupted or called away from the bedside.

In another study of the application of clinical guidelines in a large American medical center, a variation of the evaluation framework described in this paper was employed [7]. In an initial phase, usability testing was conducted involving eight primary care physicians, who were asked to “think aloud” as they interacted with an institutional electronic health record with embedded clinical guidelines. During this phase of testing, subjects were guided by the test script to interact with key features and elements of the electronic health record that included best practice advisories that were to be tested. The scripts triggered the guidelines so that users’ reactions to the guidelines could be obtained. This part of the study revealed a number of critical usability problems (e.g. such as lack of visibility of the initial alert indicating that there was a best practice advisory). After fixing these initial usability problems, a further phase of testing was conducted applying clinical simulations. This involved eight subjects who interacted with the guidelines while “interviewing” a simulated patient (which consisted of an interactive digital patient). This phase revealed issues that could not have been detected by the initial usability testing alone. This included identifying that the guidelines did not automatically trigger (i.e. be invoked) at points where the designers had expected it would. Furthermore, several subjects did not like the way the guidelines were triggered in the context of the particular style of system interaction they preferred while interacting with a patient. At the end of the clinical simulation, appropriate adjustments were made to the system based on the clinical simulation. Then a final phase involving limited testing in the live setting was conducted (i.e. with actual physician-patient-computer data collected) prior to the widespread release of the guidelines. Subsequent clinical trials indicated high uptake of the guidelines by end users (i.e. physicians in the hospital) which was attributed in part to the detection of the full range of issues discovered during the multi-phase evaluation.

A recent study which employed a multi-phase approach has indicated that there are clear benefits for conducting several layers of testing prior to system release [8]. This project involved doing several phases of testing prior to release of a regional disease management system. Data collected includ-

ed the cost of carrying out the full phases of evaluation. The benefits of conducting the evaluation were calculated and it was found that the benefits outweighed the costs by a factor of ten in the most conservative estimations. When the cost of potential technology-induced errors was considered, benefits were considerably greater. Along these lines, the clinical simulation testing phase detected errors that would have gone undetected if only the traditional system and usability testing had been carried out – i.e. the system’s body mass index calculator was found to miscalculate body mass, which could potentially lead to medication error and adverse events. It should be noted that the system would normally have been scheduled for release without the added layer of simulation testing, indicating that such testing may also provide a “safety net” that may catch errors that should have been caught much earlier (i.e. during black box system testing, or during traditional usability testing).

The approach described in this paper can also be applied to the analysis of usability problems with mobile devices and eHealth applications. In one such study, the occurrence of usability problems (obtained from conducting both usability testing and clinical simulations) was related statistically to the occurrence of technology-induced errors in a physician prescribing system. Low-cost recording methods were used in this study to record users (i.e. physicians) as they interacted with a hand-held device to enter medications [5]. This work is currently being extended to the study of the use of personal health records, mobile devices and social media used by healthcare consumers in a variety of settings, using a remote screen recording approach in both simulation and naturalistic studies. In addition, both heuristic evaluation and cognitive walkthrough can be applied in analyzing consumer e-Health software and products [39].

As a final example, in a recent study of deployment of an electronic whiteboard in hospitals in Denmark, the approach described in this paper has been used and extended to the analysis of live recordings of users interactions with the electronic whiteboard over a several month period at two hospital sites [9]. The recordings of the live interactions indicated a range of system inefficiencies, usability and workflow problems that had not been

anticipated by the designers of the electronic whiteboard. Furthermore, these issues had not been caught by any prior levels of testing that had been conducted with the electronic whiteboard. This study was conducted prior to the whiteboard going live at many locations and outlines the need for additional live testing of systems prior to widespread system release, even after usability testing and clinical simulations have been carried out.

Discussion

Ensuring the usability and safety of HIT is now recognized as being critical and important international problem [38]. New initiatives are underway in response to increased reports of technology-induced error. For example, in the United States, the Institute of Medicine (IOM) has issued a report recommending changes be made in the regulation of the safety of HIT [10]. In addition, a related report commissioned by the IOM has recommended improved methods be employed to ensure usable and safe systems [11]. In this paper, we have described a range of approaches to the analysis of usability problems and potential technology-induced errors. In addition we have described a multi-phase framework for guiding application of usability engineering methods for system safety. We have successfully employed the methods and framework in a range of studies for the evaluation of a variety of complex HIT.

From our work, we have found that it is essential to conduct user testing of HIT such as electronic health records at multiple levels, which is embodied in our evaluative framework as a number of sequential phases of testing, beginning with basic usability inspection and testing, followed by clinical simulations and finally by limited naturalistic study when possible. Just as a phased approach is taken with drug trials before the release of medications, we argue that an analogous layered “phased” approach is needed in preparing HIT for release to user populations (e.g. physicians, nurses and other health professionals). Furthermore, such optimization of HIT through phased evaluation is recommended prior to conducting traditional controlled trial studies and summative evaluation of HIT, as release

of systems that contain basic usability problems, safety or workflow issues will likely lead to low levels of adoption and will not fully explore the potential of the technology advances being studied. Furthermore, such testing needs to include the realism of clinical simulations, which the authors argue there is a need for more layers of testing other than laboratory-based usability testing. Although standard usability testing has become more commonly used by vendors of HIT for system development, subsequent clinical simulation testing is needed in the context of real use of systems (i.e. in-situ testing) in a variety of healthcare settings. We have also endeavored in our work to decrease the cost and time taken for conducting the analyses described in this paper so that combined usability testing and clinical simulations can be carried out in a rapid and cost effective manner (see [8] and [12]). Such integrated testing will be needed in order to lead to more usable and safer systems.

References

- Borycki EM, Kushniruk AW, Keay L, Kuo A. A framework for diagnosing and identifying where technology-induced errors come from. *Stud Health Technol Inform* 2009;37:56-76.
- Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. *J Biomed Inform* 2004;37: 56-76.
- Preece J, Rogers Y, Sharp H., *Interaction Design: Beyond Human-Computer Interaction*. New York: John Wiley & Sons; 2002.
- Borycki E, Kushniruk A, Kuwata S, Kannry J. Use of simulation in the study of clinician workflow. *Proc AMIA Annu Symp* 2006. p. 61-5.
- Kushniruk AW, Triola M, Borycki E, Stein B, Kannry J. Technology induced error and usability. *Int J Med Inform* 2005;74:519-26.
- Kushniruk A, Borycki E, Kuwata S, nd Kannry J. Predicting changes in workflow resulting from healthcare information System: ensuring the safety of healthcare. *Healthc Q* 2006;9:114-8.
- Li AC, Kannry JL, Kushniruk A, Chrimes D, McGinn T, Ebonyabo, et al. Integrating usability testing and think-aloud protocol analysis with "near-live" clinical simulations in evaluating clinical decision support. *Int J Med Inform* 2012 (in press).
- Baylis TB, Kushniruk A, Borycki EM. Low-cost rapid usability testing for health information systems: Is it worth the effort? *Proceedings of MIE* 2012, Pisa, Italy.
- Rasmussen R, Kushniruk A. The long and twisting path: An efficiency evaluation of an electronic whiteboard system. *Proceedings of ITCH* 2013, Victoria, Canada (in press).
- Institute of Medicine. *Health IT and patient safety: building safer systems for better care*. Committee on Patient Safety and Health Information Technology, 2011. Available from: http://www.nap.edu/catalog.php?record_id=13269
- Ash J, Kilo C, Sahpiro M, Wasserman J, McMullen C, Hersh W. Roadmap for provision of safer healthcare information systems: Preventing e-Iatrogenesis 2011, report commissioned for the Institute of Medicine.
- Kushniruk A, Borycki E. Low-cost rapid usability engineering: designing and customizing usable healthcare information systems. *Healthc Q* 2006;9(4):98-102.
- Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. *JAMA* 2005;293(10):1197-203.
- Magrabi F, Ong M-S, Runciman W, Coiera E. Using FDA reports to inform a classification for health information technology safety problems. *J Am Med Inform Assoc* 2012;19:45-53.
- Borycki EM, Kushniruk AW. Scenario-based testing of health information systems (HIS) in electronic and hybrid environments. *Stud Health Technol Inform* 2009;143:284-89.
- Yui BH, Jim WT, Chen M, Hsu JM, Liu CY, Lee TT. Evaluation of computerized physician order entry system-a satisfaction survey in Taiwan. *J Med Syst* 2012;36(6):3817-24.
- Magrabi F, Ong M-S, Runciman W, Coiera E. An analysis of computer-related patient safety incidents to inform the development of a classification. *J Am Med Inform Assoc* 2010;17:663-70.
- Borycki EM, Kushniruk AW. Where do technology-induced errors come from? Towards a model for conceptualizing and diagnosing errors caused by technology. In: *Human, Social and Organizational Aspects of Health Information Systems*. Hershey, NY: Idea Group; 2008.
- Borycki EM, Kushniruk AW, Bellwood P, Brender J. Technology-induced errors. The current use of frameworks and models from the biomedical and life science literatures. *Methods Inf Med* 2012; 51(2):95-103.
- Borycki EM, Keay, E. Methods to assess the safety of health information systems. *Healthc Q* 2010;13:47-52.
- Sittig S, Singh H. Electronic health records and national patient-safety goals. *N Engl J Med* 2012;367(19):1854-60.
- Kuwata S, Kushniruk A, Borycki E, Watanabe H. Using simulation methods to analyze and predict changes in workflow and potential problems in the use of a bar-coding medication order entry system. *AMIA Annu Symp Proc* 2006:994.
- Beuscart-Zephir MC, Pelayo S, Bernonville S. Example of a human factors engineering approach to a medication administration work system: Potential impact on patient safety. *Int J Med Inform* 2010;79(4):35-42.
- Eslami S, Abu-Hanna A, de Keizer NF. Evaluation of outpatient computerized physician order entry systems: A systematic review. *J Am Med Inform Assoc* 2007;14(4):400-6.
- IOM Roadmap. Roadmap for provision of safer healthcare information systems: Preventing e-Iatrogenesis; 2012.
- Carvalho CJ, Borycki EM, Kushniruk AW. Ensuring the safety of health information systems: Using heuristics for patient safety. *Healthc Q* 2009;12:49-54.
- Borycki EM, Kushniruk AW, Kuwata, S, Kannry J. Use of simulation approaches to the study of clinician workflow. *AMIA Annu Symp Proc* 2006:61-5.
- Borycki EM, Kushniruk AW, Anderson J, Anderson M. Designing and integrating clinical and computer-based simulations in health informatics: From real-world to virtual reality. In: Cakaj S editor. *Modeling Simulation and Optimization-Focus on Applications*; 2010. p. 31-52. Vukovar: Croatia: In-Teh. <http://www.intechopen.com/books/modeling-simulation-and-optimization-focus-on-applications/designing-and-integrating-clinical-and-computer-based-simulations-in-health-informatics-from-real-wo>
- Ammenwerth E, Hackl WO, Brinzer K, Christoffersen THE, Jensen S, Lawton K, et al. Simulation studies for the evaluation of health information technologies: Experiences and results. *HIM J* 2012;41(2):14-20.
- Nielsen J. *Usability engineering*. New York: Morgan Kaufman; 1993.
- Zhang J, Johnson TR, Patel VL, Paige D, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. *J Biomed Inform* 2003;36(1-2):23-30.
- Lewis C, Polson P, Wharton C, Rieman J. Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use. *Interfaces Chi '90 Proceedings*;1990. p. 235-42.
- Microsoft. *Microsoft Health common user interface*; 2010; <http://www.mscai.net/>
- National Health Service. *Design for patient safety: Guidelines for safe on-screen display of medication information*; 2010. Available from: <http://www.nrls.npsa.nhs.uk/resources/?entryid45=66713>
- Carvalho C, Borycki EM, Kushniruk AW, Kuwata S, Watanabe H. Simulations to assess medication administration systems. In: Staudinger B, Hoess V, Ostermann H, editors. *Nursing and Clinical Informatics: Socio-Technical Approaches*. Hershey, Pennsylvania: IGI Global; 2009. p. 144-59.
- Borycki EM, Kushniruk AW, Carvalho C. Extending heuristic evaluations to assess safety and workflow aspects of health information systems. *5th International Symposium on Human Factors Engineering in Health Informatics*. 2001, Aug 26-27. Trondheim, Norway: Tapir Academic Press.
- NHS Direct; 2012. <http://www.connectingforhealth.nhs.uk/systemsandservices/data/cui>
- Kushniruk AW, Bates DW, Bainbridge M, Househ MS, Borycki EM. National efforts to improve health information system safety in Canada, the United States and England. *Int J Med Inform*. in press.
- Monkman H, Kushniruk A. Applying usability methods to identify health literacy issues: An example using a personal health record. *Stud Health Technol Inform* 2013;183:179-85.

Correspondence to:
Elizabeth Borycki PhD
School of Health Information Science
University of Victoria
Victoria, British Columbia
Canada
Tel: +1 250 472 5432
E-mail: emb@uvic.ca