User-centered design improves the usability of drug-drug interaction alerts: Experimental comparison of interfaces

Daniel R. Luna [a,b,*], Daniel A. Rizzato Lede [a], Carlos M. Otero [a], Marcelo R. Risk [b,c],
Fernán González Bernaldo de Quirós [a]

[a] Health Informatics Department, Hospital Italiano de Buenos Aires, Argentina
[b] Instituto Tecnológico de Buenos Aires (ITBA), Argentina
[c] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

ARTICLE INFO

ABSTRACT

Clinical Decision Support Systems can alert health professionals about drug interactions when they prescribe medications. The Hospital Italiano de Buenos Aires in Argentina developed an electronic health record with drug-drug interaction alerts, using traditional software engineering techniques and requirements. Despite enhancing the drug-drug interaction knowledge database, the alert override rate of this system was very high. We redesigned the alert system using user-centered design (UCD) and participatory design techniques to enhance the drug-drug interaction alert interface. This paper describes the methodology of our UCD. We used crossover method with realistic, clinical vignettes to compare usability of the standard and new software versions in terms of efficiency, effectiveness, and user satisfaction. Our study showed that, compared to the traditional alert system, the UCD alert system was more efficient (alerts faster resolution), more effective (tasks completed with fewer errors), and more satisfying. These results indicate that UCD techniques that follow ISO 9241-210 can generate more usable alerts than traditional design.

1. Introduction

Recent research has highlighted that medical errors are the third leading cause of death in the United States, following heart diseases and cancer [1]. In 1999, the Institute of Medicine reported that almost 25% of total medical errors involved prescription medication [2]. Although most errors are harmless, a few can cause severe injuries and even death. Leape et al. found that half of these errors occur when the medication is first prescribed [3]. Sometimes, errors occur when the prescriber does not know about relevant drug-drug interactions (DDIs) [4,5]. Adverse drug events related to DDI are generally preventable, but Glassman et al. have noted that clinicians do not recognize them at least half of the time [6].

Increasingly popular methods to help order providers on a large scale include clinical decision support systems (CDSS). Clinical decision support is defined as "the use of information and communication technologies to bring relevant knowledge to bear on the healthcare and well-being of a patient" [7]. These systems have been shown to improve both quality of care and resource optimization [8–10]. Although electronic prescription systems with real-time detection of DDIs appear to be the most suitable implementations, several studies showed that they still have low performance and high alert override rates [11–15]. As reported in the scientific literature, these systems have at least five potential drawbacks:

1. Inaccurate knowledge databases trigger excessive alerts (mostly with low clinical significance), leading to alert fatigue [16,17].
2. Rudimentary interfaces lack intuitive design and workflow integration [18].
3. The systems lack relevant medical context data, and thus their rules can be too simplistic [19,20].
4. Low monitoring levels hinder continual improvement processes [21].

5. Implementation varies due to lack of standards even for the same vendor [22].

In the mid-2000s, the Hospital Italiano de Buenos Aires, Argentina, used traditional software engineering to develop an in-house electronic health record (EHR) system with a computerized physician order entry (CPOE) and DDI alert system. Clinical pharmacology specialists reviewed different sources, including textbooks, product monographs and relevant literature to create the knowledge database. This database was periodically updated using publications on pharmacological interactions like "Evaluation of Drug Interactions", a former report by First Data Bank Inc. [23]. The uploaded data was referenced to SNOMED CT, the terminology standard for health information contained in the patient clinical data repository [24]. After some time, we evaluated the performance of our clinical decision support system and found a high alert override rate. Thus, we focused on enhancing the quality of the knowledge database, and then published our experience [25]. Our analysis included systematic evaluation of each DDI according to clinical relevance [26], to eliminate combinations with a low probability of harm (false positives), as suggested in a recent consensus [27]. We also created alert severity tiers, according to potential reaction seriousness, following recommendations from Paterno et al. [17]. As alert acceptance remained persistently low after these changes, the chief medical information officer withdrew the DDI decision support system to identify other potential issues. Previous research by Seidling et al. found that the alert display quality most strongly predicted DDI alert acceptance [28]. Therefore, we endeavored to improve the alert design by applying user-centered design (UCD) techniques, as they have demonstrated to increase adoption and usage efficiency of health information technology tools [29,30].

According to Patel and Kannampallil, human-computer interaction is a fundamental aspect to consider when developing computer systems [31,32]. UCD is a process framework that makes a system usable and understandable by accounting for end-users' needs, wants and constraints, through the whole product cycle. Following perspectives from Norman, UCD starts by understanding and specifying the context and requirement analysis, and then designing and iteratively testing solutions [33]. This systematic process is regulated by ISO 9241-210 "Human-centered design for interactive systems" [34]. For Kushniruk et al., participatory design goes beyond UCD and cooperative design approaches to include end users as active participants in the design and decision making [35]. Usable systems have many benefits including enhanced productivity, error reduction, reduced training costs and increased acceptance [36].

Although in recent years the quality of software engineering studies has improved gradually, most of them lack an experimental design and statistical methods [37,38]. In a survey that included more than 5000 research papers from 1993 to 2002, just 1.9% were controlled trials [39]. Marcilly et al. recommend that evidence-based usability engineering should focus on design elements and especially on usability evaluation methods [40].

This paper aims to describe the methodology of our UCD, which includes participatory design, to revise the drug-drug interaction alert interface and to test a crossover method for scientifically comparing the usability of standard and new interface in terms of efficiency, effectiveness and user satisfaction.

2. Methods

2.1. Setting

The Hospital Italiano de Buenos Aires (HIBA) is a non-profit healthcare academic center founded in 1853, with more than 2700 physicians, 2700 other health team members (including 1200 nurses), and 1800 administrative and support employees. The HIBA network includes two hospitals in Buenos Aires city and its suburban area, 750 beds (200 for intensive care), 41 operating rooms, 800 home care beds, 25 outpatient clinics and 150 associated private practices. It has a Health Maintenance Organization (Plan de Salud) that covers more than 150,000 people and provides health services to another 1,500,000 people who are covered by affiliated insurers. Between 2013 and 2014, the HIBA admitted more than 45,000 inpatients, conducted 45,000 surgical procedures (50% ambulatory) and 3,000,000 outpatient visits. The HIBA also is a teaching hospital, with more than 30 medical residency-training programs, 34 fellowship programs and 400 residents and fellows in training.

Since 1998, the HIBA has run an in-house developed health information system, which includes clinical and administrative data [13]. Its EHR system called Italica, is an integrated, modular, problem oriented and patient centered system that works in different clinical settings (outpatient, inpatient, emergency and home care). Italica allows computerized physician order entry for medications and medical tests, and storage and retrieval of tests results, including archived images. It was the first hospital in Argentina and the second in Latin America to be certified by the HIMSS as level 6+ in the Electronic Medical Record Adoption Model [41,42].

In recent years, our Health Informatics Department at the HIBA prioritized UCD in the design and development culture to enhance the usability of healthcare software. We conducted lectures, launched a pilot project, and assembled a usability team for service and dissemination [43]. Regarding this study, all sessions were held in the new usability laboratory. Morae® 3.2 software [44] was used for audio and video recording of the meetings.

2.2. Methodological design

Our study was divided into two different phases: the first phase involved the iterative UCD process of developing a new alert interface, and the second phase involved the experimental approach of testing the new UCD interface against the standard version, which was conducted in a controlled environment and used the same DDI knowledge database.

2.2.1. First phase: User-centered design with participatory design

We reformed the DDI alerts by using a Participatory Design approach based on UCD techniques. As described in a state-of-the-art reference handbooks on the subject, Participatory Design can be defined as "a process of investigating, understanding, reflecting upon, establishing, developing and supporting mutual learning between multiple participants in collective 'reflection-in-action', where the participants typically undertake the two principal roles of users and designers where the designers strive to learn the realities of the users' situation while the users strive to articulate their desired aims and learn appropriate technological means to obtain them" [45]. A team of three health informatics specialists and two usability experts worked with users following the ISO 9241-210. This phase took place at HIBA from September 2013 to April 2014, and was undertaken in three stages (inquiry, participatory design, and usability testing), as described in a previous publication [46]. The users were physicians that worked in outpatient and inpatient settings and had at least four years of experience with the CPOE system. Patient scenarios (clinical vignettes) were developed based on real clinical cases [47,48], taking the most frequent and significant examples of DDI from the clinical data repository [14].

2.2.1.1. Stage 1: Inquiry. Stage 1 involved interviews and contextual observations [49] of the electronic prescribing process in situations

with potential drug interactions. A semi-structured questionnaire was given to the users to guide them through the dynamic clinical scenarios. At this stage, the users interacted with the standard DDI alert version. Low-fidelity prototypes of new DDI alerts interfaces were generated using pen and paper sketches, accounting for ideas that emerged from the interviews.

2.2.1.2. Stage 2: Participatory design. Stage 2 started with low-fidelity prototypes. Two different prototypes made in Balsamiq® [50] were used: one was printed on paper and the other was displayed as an interactive electronic presentation to reproduce the prescription process as closely as possible. We shared the clinical vignette and asked the users to prescribe certain medications, leading to potential DDIs. When the alert popped up, we asked them to solve it. Physicians' opinions and feelings were obtained and recorded using the Think Aloud technique [51]. The approach was qualitative, pursuing domain saturation. The participating physicians guided the development of new prototypes, giving constant feedback. Two cycles of prototyping and testing were conducted.

2.2.1.3. Stage 3: Usability testing. For the third stage, high fidelity prototypes were developed in Axure® [52], and presented to a new group of users. The sequence was similar to the previous stage, but focusing on quantitative metrics (effectiveness, efficiency and user satisfaction).

Effectiveness was assessed by the override rate and the task completion rate.

Efficiency was measured as the number of clicks and time spent to finish the task.

User satisfaction was assessed using a System Usability Scale (SUS) questionnaire [53]. SUS is a Likert scale-based survey, in which a carefully selected statement is made and the respondent then indicates the degree of agreement or disagreement with it on a 5-point scale [54].

The process was iterative; each stage included prototyping cycles for domain saturation to reach the best possible model (see Fig. 1). The last UCD prototype was developed as a new software version.

2.2.2. Second phase: Experimental test of interfaces

Prior to our interventional study, we obtained Institutional Review Board approval. Users gave their informed consents to participate in the research. The study had an experimental crossover design in which each participant was twice exposed to randomly assigned clinical vignettes, working once with each interface.

We compared two different DDI alert interfaces: the standard one (developed under traditional techniques) and the final participatory design model, generated using UCD techniques. The clinical decision support system ran every time a new prescription was placed and searched the knowledge database for potential

Fig. 1. The iterative process of the alert UCD consisted of three stages: inquiry, participatory design and prototype. The inquiry stage involved interviews and contextual observations of the drug-prescribing process in situations with potential drug interactions. The design stage used two cycles of low-fidelity prototypes. Finally, the prototype stage was performed with high-fidelity prototypes.

interactions between each substance already on the list and the new drug. When an interaction was detected, the system opened a DDI alert modal (a dialog box or popup window that was displayed on top of the current page), in the standard and UCD versions.

The standard DDI alert interface had different elements (see Fig. 2): the name of the causative drugs; the clinical significance (i.e., severity tier); a short explanation of the interaction that could be expanded through a "learn more" button; and two action buttons (i.e., "Cancel" or "Ignore"). Canceling the alert implied stopping the prescription of the second drug whereas ignoring the alert kept both drugs. The alert was displayed using colored boxes, and the recommendation appeared as text for the physician to read.

The new clinical decision support system shared the same DDI knowledge database and inference engine but diverged in the alert interface (see the First Phase Results section for a description). All participants were individually exposed to both interfaces in two different sessions, with one month period in between. The sequence of the interface type exposition was randomized (old-new; new-old) among participants.

Medical experts used frequent DDI examples extracted from a local yearly prescription database [14] to develop 12 clinical vignettes [47,48] (four for each clinical setting: outpatient, non-critical inpatient and critical inpatient). Therefore, every scenario was constructed from a single significant DDI and presented under carefully controlled situations to reflect reality. The sessions, which were held in our Usability Lab with one user and two researchers, started with a brief introduction followed by a short interview with the user. During the session, we delivered four randomly selected clinical vignettes and the test instructions. In each clinical case the participant had to follow the directives regarding electronic prescription of certain drugs and take actions within the system when it showed the DDI alert. Participants were asked to share their thoughts in real time while doing the task (Think Aloud) [51]. Direct observation and recording were performed. At the end of the meeting, every participant answered a semi-structured survey about the alerts.

2.2.2.1. Study population. Physicians fulfilling the exclusion and inclusion criteria were randomly selected for each clinical setting (outpatient, non-critical, and critical inpatient). All had worked with CPOE but had not been recently exposed to the original DDI alert, because it had been withdrawn more than a year prior to the study.

Inclusion criteria were as follows:

- Working as a physician of any specialty in the Hospital for more than a year by March 2012.
- Working in the selected clinical setting (outpatient, critical, or non-critical inpatient) for more than six months.
- Having a number of interactions of ±1 standard deviation of the mean in the previously performed DDI database analysis [14].

Exclusion criteria were as follows:

- Study participation refusal.
- Previous participation in UCD stages.

2.2.2.2. Sample calculation. The sample was calculated with Power and Precision software version 3.2.0. It was necessary to include 30 physicians (distributed in 10 physicians *per stratum*) to test the null hypothesis (i.e., no difference in the percentage of failed prescriptions between the two interfaces), with an expected difference of at least between 45% and 20%, a type 1 error of 5%, and an estimated power of 80%.
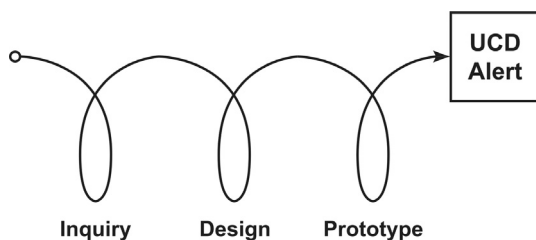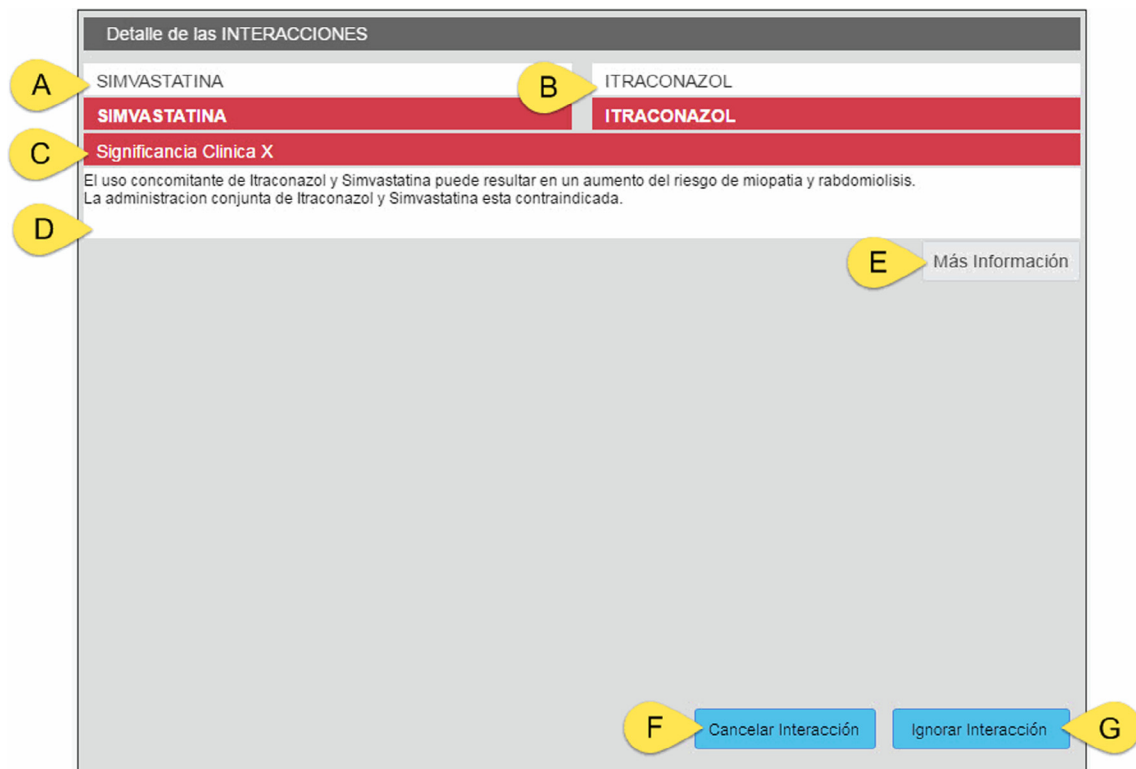
Fig. 2. Standard DDI alert interface (letters correspond to descriptions): the two drugs (A and B), the clinical significance (C), a brief explanation (D), a "learn more" button (E), and the action buttons to Cancel the prescription (F) or Ignore the alert (G).

2.2.2.3. Measurements. Efficiency, effectiveness and user satisfaction were selected as usability metrics to compare the performance of the new UCD alert against the standard version.

Efficiency was measured through three variables extracted from Morae® recordings: time spent in solving the alert, number of mouse clicks to complete the task, and number of words used in the text entry.

To assess effectiveness, in each clinical case the user was rated when solving the alert as successful, successful with problems, or unsuccessful. The laboratory test allowed us to measure the errors and error tolerance for each user interface. The analysis of the recordings allowed comparison of the number and type of errors and the error recovery.

At the end of the session, users were invited to complete a satisfaction questionnaire based on the SUS [53,54]. A brief semi-structured interview, based on a survey developed by Zheng et al. [55], was conducted to collect user perceptions regarding the advantages and challenges of using the system.

2.2.2.4. Statistical analysis. Descriptive statistics were generated for all variables. Interval variables were parameterized by mean and standard deviation. For categorical variables, the observed frequency (total number of observations within the category) and relative frequency percentages were used. Statistical analyses for all tests were performed using the R software environment from R Project for Statistical Computing [56]. Statistical significance was considered when the probability was lower than 0.05.

Efficiency was statistically analyzed by three-pathway Analysis of Variance (ANOVA). Each of the three variables of efficiency (Time, number of Clicks and number of Words) was a response (dependent variable), and the three independent variables were the three pathways (traditional and UCD Design Methodology; Outpatient, Critical Inpatient, and Non-Critical Inpatient Clinical Setting; and User). The ANOVA was performed to test the null

hypothesis that "efficiency variables are not different regarding the design methodology". Clinical Setting and User were control variables. Interaction graphs for each ANOVA were performed.

Effectiveness was statistically analyzed using the Chi square test for a 2-by-3 table and Fisher's exact test for a 2-by-2 table for categorical variables. The odds ratio was calculated to estimate the association between design methodologies and the proportion of complete responses with or without serious errors and incomplete responses (English odds ratio, with confidence interval of 95%).

For user satisfaction analysis as a percentage, the Wilcoxon signed-rank test for repeated measurements was used. Each user was measured twice, one for each interface design.

3. Results

3.1. First phase: User-centered design with participatory design

Twenty-four physicians from three different fields (outpatient, critical inpatient, and non-critical inpatient) participated in the three stages.

3.1.1. Stage 1: Inquiry

Six physicians participated in the first stage (Inquiry). Initial meetings and interviews provided feedback on the standard interface (see Fig. 2). Participants made several comments, some of which have already been described in the literature. Here, we quote some illustrative commentaries.

Regarding their experience on DDI information retrieval from other sources, one person responded,

"It is very hard for me to find relevant information about drug-drug interactions when prescribing. If I look for them in my smartphone, almost every combination of drugs would have a potential

interaction, and I would dismiss the majority because of their null significance."

Another person shared an opinion about the standard interface's visual display,

"I think the standard alert display has a poor interface design, and it has a lot of information to read. I need a quick and useful alert to prevent potential damage to my patients, without spending hours in front of the computer."

Concerning a previous experience with the DDI alert, one person noted,

"I was tired of canceling unnecessary drug-drug interaction alerts. I'm happy about the old decision support system withdrawal and I expect a better second version."

Regarding the available option buttons to solve the standard alert, one person stated,

"If I have to exit the alert and then restart the prescription process following the recommendations, there is a chance for me not to do as suggested."

This feedback was incorporated into the structure of the first prototypes, starting with pen and paper sketches and then Balsamiq printed models (see Fig. 3).

3.1.2. Stage 2: Participatory design

At stage 2, a different sample of physicians was exposed to the printed versions (Fig. 3) of the first prototypes, the same as during the DDI alert in the clinical vignette. While interacting with the initial paper prototype, the facilitator showed subsequent elements upon request (like "learn more" information). Participants decided whether to use labels and could ask for new ones. Afterwards, users' views and recommendations were iteratively used with participatory design techniques (participatory prototyping and sketching) to improve the model that would be used in the following step. Fig. 4 shows a screenshot of the interactive electronic presentation.

The sessions were analyzed with qualitative techniques to structure their outcomes.

We solicited the following feedback about the second prototype's visual display,

"I like the different colors of the alert; when everything is red I tend to misinterpret it and it is far more difficult for me to find the proper solution."

On the subject of the available option buttons to solve the second prototype alert, one person remarked,

Fig. 3. Example of the first prototype with UCD. The opinions of six physicians were used to shape information about the DDI and the recommendation.

Fig. 4. Example of the second prototype with UCD, showing the result of the iterative design process and comprehensive display, with different colors for buttons and messages.

Fig. 5. Example of final UCD refined prototype.

"The recommendations shown as action buttons eased our lives. I think this system adapts pretty much to our workflow."

There were also negative observations,

"I find it quite complex to have so many options just for a single alert."

This feedback was incorporated into the redesign and retesting until users considered them appropriate to build the final UCD prototypes (see Fig. 5).

3.1.3. Stage 3: Usability testing

The usability testing of the prototypes focused on effectiveness, efficiency and user satisfaction. We tested the resulting prototypes of every stage.

Regarding effectiveness, 11 of 24 physicians ignored the warning (override rate of 45.8%), and 13 (54.2%) took action. Every participant finished the tasks (task completion rate of 100%): 13 physicians (54.2%) without difficulty, 10 (41.7%) with questions or minor errors, and 1 (4.2%) with serious errors.

Table 1
Efficiency (number of clicks) for each stage of phase 1.

Action	Performance	Number of clicks		
		Stage 1	Stage 2	Stage 3
Prescribing process	Completed the task	68	70	65
	Completed with questions or minor errors	71	68	65
	Completed with serious errors	–	–	–
Follow the recommended action	Completed the task	54	60	50
	Completed with questions or minor errors	52	75	70
	Completed with serious errors	–	–	–
Prescription cancelled	Completed the task	13	10	8
	Completed with questions or minor errors	76	75	70
	Completed with serious errors	–	–	–

Table 2
Efficiency (time spent) for each stage of phase 1.

Action	Time (s)		
	Stage 1	Stage 2	Stage 3
Prescribing process	16	18	15
Follow the recommended action	11	11	10
Prescription canceled	9	12	10

Table 3
User satisfaction with SUS scores for phase 1.

By question	Score
I would like to use this system when prescribing	3.0
The system is unnecessarily complex	3.0
The system is easy to use	2.9
I would need the assistance of an expert to use the system	3.9
The diverse possibilities of the system are well integrated	3.6
There is too much inconsistency in the system	3.4
I believe that most of the clinicians could use the system without problems	2.0
The system is very uncomfortable	3.1
I felt confident using the system	3.1
I would need a lot of instructions to manage the system	3.1

Efficiency was measured as the number of clicks and time needed to finish the task. Tables 1 and 2 summarize the results for the prototypes of each stage.

User satisfaction measured by the SUS gave an average value of 77.9 per participant on a scale of 0–100 (see Table 3).

3.1.4. Revised DDI interface

The final prototype was developed as a new software version. Compared to the standard version (Fig. 2), it had a different communication message, changing the displayed elements, warning colors and proposed actions (as shown in Fig. 6). New output actions were specifically created for the novel DDI alerts.

3.2. Second Phase: Experimental test of interfaces

Out of 191 physicians (potential candidates), 150 (50 of each clinical setting: outpatient, critical inpatient and noncritical inpatient) were invited to participate in the study. Then, 10 participants from each stratum (i.e., 30 physicians) were randomly selected and included in the study. Three of them (two participants from critical inpatient and one from outpatient) did not complete the second stage of evaluation. Thus, new physicians were recruited to include 10 participants for each clinical setting. Table 4 shows the demographics for each variable and group.

The efficiency analysis was performed with the statistical model, *Dependent Variable ~ Design Methodology + Clinical Setting + User*. "Time spent by users", "Clicks", and "Words" were the dependent variables. "Design Methodology", "Clinical Setting", and "User" were the pathways. Table 5 shows the results of ANOVA for each dependent variable and pathway. Figs. 7–9 show the interaction plots for each ANOVA. Regarding the design methodology for inpatients, "Time spent by users" was less for the UCD version (p = 0.0004) but there were no significant differences for "number of Clicks" and "number of Words".

Effectiveness was studied using the reports and recordings about whether the tasks were incomplete or completed with
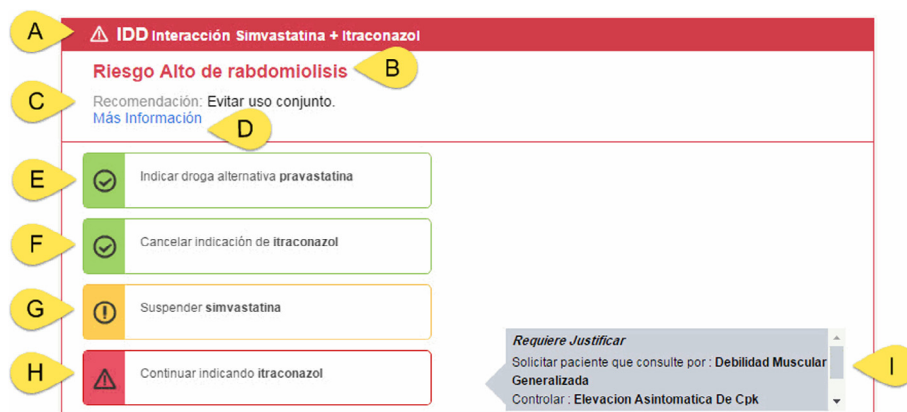

Fig. 6. New software version interface (letters correspond to descriptions) displaying the drugs involved (A), the drug interaction risk and severity (B); the main recommendation (C); a "learn more" link (D); the action buttons to "prescribe an alternative drug" (E), "cancel drug 2 prescription" (F), "cancel drug 1 prescription" (G) and the "keep both drugs" option to ignore the alert (H), which required justification and monitoring as stated in the text bubble (I).

Table 4
Demographics for each variable and group.

	Total (N = 30)	Outpatient (10)	Critical Inpatient (10)	Non-critical Inpatient (10)	p
Age[a] in years	34.3 (3.4)	34.2 (3.4)	36 (1.2)	32.7 (4.1)	0.084
Female%[b]	56.7 (17)	60 (6)	50 (5)	60 (6)	0.873
Graduated[a] in years	9.6 (2)	9.5 (2.2)	10.6 (0.8)	8.3 (2.3)	0.079
EHR use[a] in years	6.7 (0.9)	6.8 (1.1)	6.6 (0.8)	6.6 (0.8)	0.864
Time between evaluations[a] in days	16 (2.6)	17.1 (3.4)	15.7 (1.9)	15.7 (2.1)	0.245

[a] Average (standard deviation).
[b] Relative frequency in percentage (amount).

Table 5
Efficiency analysis: probability table of three-pathway ANOVA.

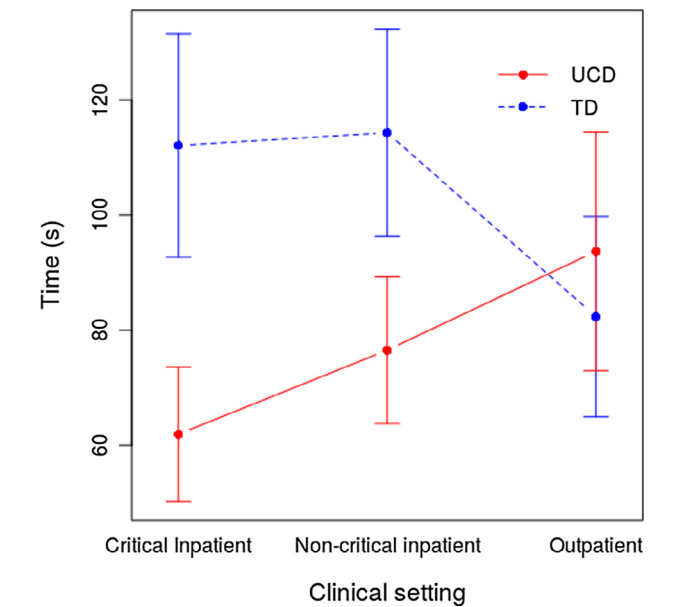| ANOVA pathway	p for Time	p for Clicks	p for Words
Design methodology	0.0004	0.69	0.094
Clinical Setting	0.58	0.15	<0.0001
User	0.59	0.70	0.21

Fig. 7. Efficiency interaction plots by "Time spent by users" for the traditional design (TD) and UCD methodology interfaces. These plots include the dependent variables in the y-axis. The x-axis factor is the Clinical Setting. The trace factor is the Design Methodology. The points show the mean value, and the error lines represent ±2 standard error.

Fig. 8. Efficiency interaction plots by "number of mouse Clicks" used to finish the task for the traditional design (TD) and UCD methodology interfaces. These plots include the dependent variables in the y-axis. The x-axis factor is the Clinical Setting. The trace factor is the Design Methodology. The points show the mean value, and error lines represent ±2 standard error.

Fig. 9. Efficiency interaction plots by "number of Words" written for traditional design (TD) and UCD methodology interface. These plots include the dependent variables in the y-axis. The x-axis factor is the Design Methodology. The points show the mean value, and error lines represent ±2 standard error.

severe errors, with minor errors or doubts, and with no errors. Table 6 is a 2-by-3 table, in which the columns show the interface types (traditional method; UCD) and the rows rank the task fulfillment of the reports. The chi-square test of independence showed a chi-square statistic = 5.79, with 2 degrees of freedom, and a p-value = 0.055. Prima facie, it means that there were no significant differences in global effectiveness of the design methodologies.

As responses with minor errors or doubts had the same absolute frequency for both design methodologies (12), the corresponding row was excluded and a new 2-by-2 table was then created, as shown in Table 7. The analysis of this sub-table with Fisher's exact test for 2-by-2 tables showed a p = 0.045, with a true odds ratio of 10.2 (95% CI; 1.1 to 514). I.e., users were 10.2 times more likely to solve the DDI alerts without errors when using the UCD interface. This result confirms differences in the effectiveness of the designs: the traditional design produced more reports of errors than the UCD, and the UCD resulted in more reports of error-free accomplished tasks.
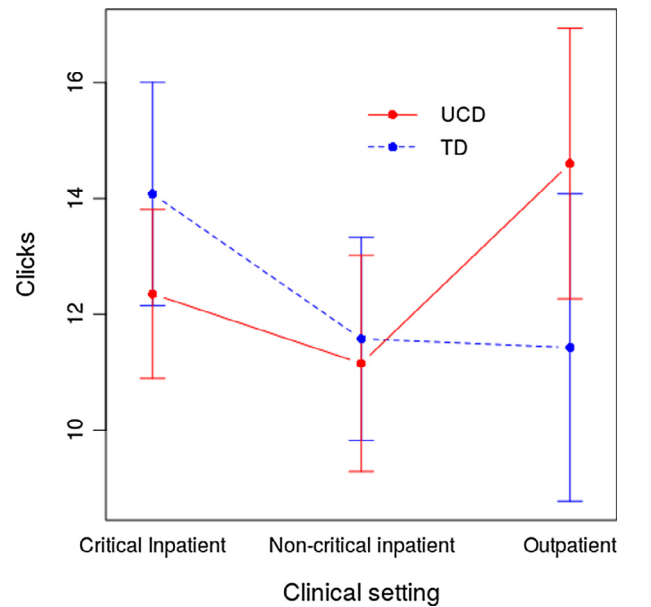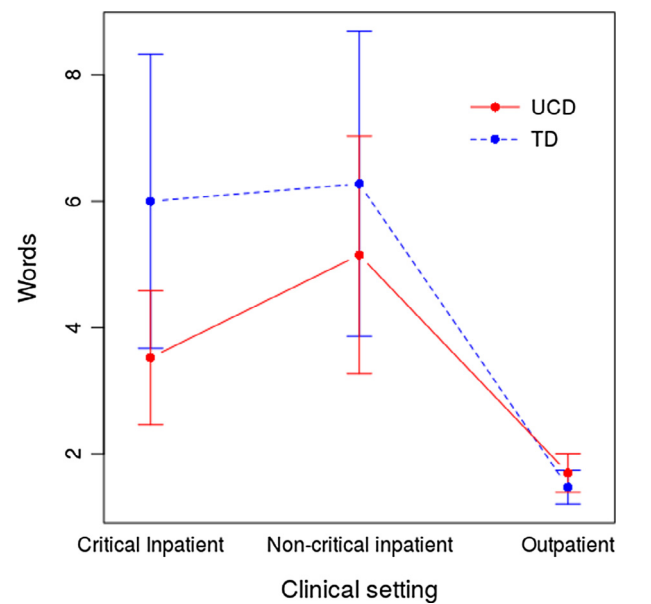
Table 6
Effectiveness analysis between the interface designed with the traditional method and that designed with UCD methodology.

DDI reports	Traditional design	UCD
Incomplete performance or complete with severe errors	7 (23)	1 (3)
Complete with minor errors or doubts	12 (40)	12 (40)
Complete with no errors	11 (37)	17 (57)

Table 7
Effectiveness analysis between the interface designed with the traditional method and that designed with UCD methodology, excluding minor errors.

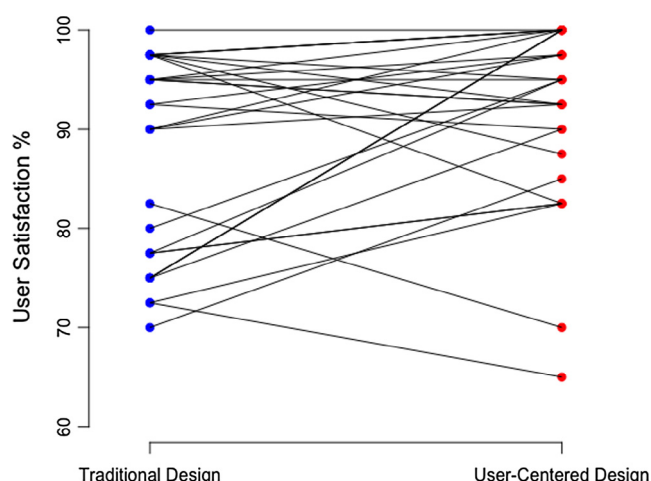| DDI reports	Traditional design	UCD
Incomplete performance or complete with severe errors	7 (39)	1 (6)
Complete without errors and mistakes	11 (61)	17 (94)

Fig. 10. Analysis of user satisfaction, comparing the traditional design and UCD. Satisfaction for each user is represented by dots connected with lines. The average user satisfaction for the traditional design was 87.4% and for the UCD it was 92% (p = 0.024).

For the analysis of user satisfaction when using the interfaces, a test with repeated measures was performed, using the SUS data. For the traditional design the average user satisfaction was 87.4%, whereas for the UCD methodology it was 92%, p = 0.024 (see Fig. 10). Out of the 30 participants of the study, 19 (59.3%) felt more satisfied when using the UCD interface than when using the standard interface. Two participants (7.4%) found no differences in satisfaction using either interfaces, and nine users (33.3%) preferred the standard design.

4. Discussion

Regarding the first phase, the initial interviews gave us insights about users' needs, tasks to accomplish and desired functionalities. Based on that information, usability designers sketched the first low-fidelity prototypes. Most physicians were concerned about DDI and expressed the need for decision guidance when solving the alert.

Our study included two cycles of participatory design sessions and prototypes and evaluations focused on qualitative aspects. Participants agreed that they wanted short, clear, and quick alerts, as described in previous research [18,19]. They also complained about the difficulty in finding relevant drug interaction information and

about the poor interface design, which supported the hypothesis of usability flaws recently described by Marcilly et al. [57]. We incorporated this feedback into our subsequent prototypes to improve the presentation of the information and to simplify the reasoning process. The workflow integration was the most complex issue in the interface design. An important outcome of this stage was the requirement of an action-oriented alert, or "the ability to take actions within the alert box, without interrupting the workflow or restarting the prescription process" [58]. Integrated, one-click actions (e.g., "change dose", "cancel first drug", "cancel second drug", "keep both prescriptions") were perceived as a great advantage.

The Usability Testing of the prototypes showed an upward trend in effectiveness, efficiency and user satisfaction. To validate these results we performed an experimental evaluation (second phase).

The second phase showed that our participatory design method was a reliable way of designing and developing DDI alerts. Our results are similar to those in Russ et al. [30]. Regarding efficiency the new interface required less time to complete the task but the same number of clicks and justification words as in the old one. These findings may indicate a sign of quick and enhanced interaction with the alert [30].

The UCD interface showed statistically significant improvements in effectiveness (as measured by the number of cases completed without errors).

User satisfaction was also better with UCD methodology. The analysis of user preference showed that nearly 60% preferred the UCD interface. This finding was evident in interviews, in which the main findings highlighted two fundamental concepts: the relevant option on the interaction was always visible and the recommended action was available on the same screen. The users also found the color and orientation as positive changes and preferred having more than two options, rather than just "ignore" or "continue". Another interesting result was that, even with more information on the screen, they perceived that the new interface allowed faster resolutions than the old one. A key advantage highlighted by the users was the action-oriented aim of the new interface or "the ability to take actions within the alert box, without interrupting the workflow or restarting the prescription process", as suggested in previous studies [22,58].

The main highlights of the UCD interface included color as a risk indicator; an action-oriented aim; and relevant, short, and accessible information as recommended by literature [19,57–61]. These factors might be responsible for the improvements in usability metrics.

In our research, the acceptance and override rate differed significantly depending on the clinical setting. Potential for drug-drug interactions increases as the number of drugs increases (polypharmacy), as frequently seen in complex, critical inpatients cares [62–64]. Therefore, the number of omissions can increase because benefits outweigh risks in these situations [65]. Users thus appreciated the follow-up recommendations to monitor potential adverse events. Additionally, combinations of interacting drugs are sometimes used intentionally with favorable effects. In the intensive care setting, such combinations are commonplace for sedation, analgesia, and other supportive care. Consequently, alerts are overridden in these circumstances, as they are not seen as a special risk [66]. On the contrary, infrequent high-risk interactions in outpatient settings can promote the acceptance of such DDI alerts.

Regarding flaws in the participatory design approach, we had difficulties recruiting participants. Although the HIBA has nearly 3000 physicians, their accessibility is low. Thus, we recommend using a budget to cover professional research time and avoid conflicting with office hours. It was also difficult to create clinical

vignettes that were understandable and realistic but also easy to solve. This should be considered to avoid clinician misunderstanding in the efficiency and effectiveness measures.

We followed Kushniruk et al.'s [35] suggestion about using portable devices to conduct this kind of study in the users' environments. This way, participants did not need to visit a single physical location to perform the tests.

The topic of this paper has been studied for many years [32,67,68], with varying results. The main contribution of this study is the proof of concept of a rigorous scientific method for measuring and comparing efficiency, as well as effectiveness of and user satisfaction with a UCD interface.

4.1. Limitations

The usability testing was done in a laboratory under carefully control to reflect reality. Thus, it was not a real-world evaluation, in which things can change dramatically in unpredictable ways. Additionally, our research was done in a single academic center using in-house developed software and thus might not represent other institutions.

4.2. Future directions

To extend this research into real environments, we are conducting the same study in a prospective way with true clinical cases. A randomized controlled trial with two branches (UCD versus traditional design) to test web-based EHR with CPOE-integrated alerts is ongoing. We expect to have preliminary results soon.

5. Conclusion

Incorporating UCD techniques into the development of support tools for DDI alerts improved usability in terms of effectiveness, efficiency, and user satisfaction, compared with the traditional design interface. However, given the multiplicity of factors that influence healthcare, such as the setting and feasibility of development, our conclusions should be carefully analyzed before extrapolating them to other scenarios. The participatory design approach enabled the usability and development teams to work with end users to understand the tasks and complexity of the process and to improve the software quality. We conclude that UCD that follows ISO 9241-210 could be applied to generate more usable alerts than traditional design.

Acknowledgments

References

[1] M.A. Makary, M. Daniel, Medical error — the third leading cause of death in the US, BMJ 2139 (2016) 1–5, http://dx.doi.org/10.1136/bmj.i2139.

[2] L.T. Kohn, J.M. Corrigan, M.S. Donaldson, To err is human: building a safer health system, National Academies Press, Institute of Medicine, 1999.

[3] L.L. Leape, D.W. Bates, D.J. Cullen, J. Cooper, H.J. Demonaco, T. Gallivan, et al., Systems analysis of adverse drug events. ADE Prevention Study Group, JAMA 274 (1995) 35–43.

[4] T.J. Bright, A. Wong, R. Dhurjati, E. Bristow, L. Bastian, R.R. Coeytaux, et al., Effect of clinical decision-support systems: a systematic review, Ann Intern Med 157 (2012) 29–43, http://dx.doi.org/10.7326/0003-4819-157-1-201207030-00450.

[5] G.J. Kuperman, A. Bobb, T.H. Payne, A.J. Avery, T.K. Gandhi, G. Burns, et al., Medication-related clinical decision support in computerized provider order entry systems: a review, J Am Med Informatics Assoc 14 (2007) 29–40, http://dx.doi.org/10.1197/jamia.M2170.

[6] P.a. Glassman, B. Simon, P. Belperio, A. Lanto, Improving recognition of drug interactions: benefits and barriers to using automated drug alerts, Med Care 40 (2002) 1161–1171, http://dx.doi.org/10.1097/01.MLR.0000036429.31445.E0.

[7] R. Greenes, Clinical decision support: the road ahead, Academic Press, 2014, http://dx.doi.org/10.1016/B978-0-12-398476-0.00029-4.

[8] B. Chaudhry, J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, et al., Systematic review: impact of health information technology on quality, efficiency, and costs of medical care, Ann Intern Med 144 (2006) 742–752.

[9] J.J. Coleman, H. van der Sijs, W.E. Haefeli, S.P. Slight, S.E. McDowell, H.M. Seidling, et al., On the alert: future priorities for alerts in clinical decision support for computerized physician order entry identified from a European workshop, BMC Med Inform Decis Mak 13 (2013) 111, http://dx.doi.org/10.1186/1472-6947-13-111.

[10] M. Xie, M.B. Weinger, W.M. Gregg, K.B. Johnson, Presenting multiple drug alerts in an ambulatory electronic prescribing system, Appl Clin Inform 5 (2014) 334–348, http://dx.doi.org/10.4338/ACI-2013-10-RA-0092.

[11] H. van der Sijs, J. Aarts, A. Vulto, M. Berg, Overriding of drug safety alerts in computerized physician order entry, J Am Med Informatics Assoc 13 (2006) 138–147, http://dx.doi.org/10.1197/jamia.M1809.

[12] Schpilberg M, Bernaldo de Quiros FG, Luna D, Gomez A, Martinez M, Cifarelli G. Creación de un sistema para la detección de interacciones farmacológicas en una Historia Clínica Electrónica. 4to Simp. Informática en Salud - 30 JAIIO, vol. 30; 2001. p. 105–13.

[13] D. Luna, F. Plazzotta, C. Otero, F. González Bernaldo de Quirós, A. Baum, S. Benítez, Incorporación de tecnologías de la información y de las comunicaciones en el Hospital Italiano de Buenos Aires, CEPAL (2012).

[14] Luna D, Waldhorn M, Ianicelli A, Risk M, González Bernaldo De Quirós F. Drug-drug interaction in a healthcare network of Argentina: frequency and characteristics. XIII Congr. Bras. em Informática em Saúde – CBIS 2012, 2012.

[15] Institute of Medicine (IOM), Health IT and patient safety: building safer systems for better care, The National Academies Press (US), Washington (DC), 2012, http://dx.doi.org/10.1001/jama.308.21.2282-a.

[16] A. Patapovas, H. Dormann, B. Sedlmayr, M. Kirchner, A. Sonst, F. Müller, et al., Medication safety and knowledge-based functions: a stepwise approach against information overload, Br J Clin Pharmacol 76 (2013) 14, http://dx.doi.org/10.1111/bcp.12190.

[17] M.D. Paterno, S.M. Maviglia, P.N. Gorman, D.L. Seger, E. Yoshida, A.C. Seger, et al., Tiering drug-drug interaction alerts by severity increases compliance rates, J Am Med Informatics Assoc 16 (2009) 40–46, http://dx.doi.org/10.1197/jamia.M2808.

[18] A. Feldstein, S.R. Simon, J. Schneider, M. Krall, D. Laferriere, D.H. Smith, et al., How to design computerized alerts to safe prescribing practices, Jt Comm J Qual Saf 30 (2004) 602–613.

[19] J.D. Duke, X. Li, P. Dexter, Adherence to drug–drug interaction alerts in high-risk patients: a trial of context-enhanced alerting, J Am Med Informatics Assoc 20 (2013) 494–498, http://dx.doi.org/10.1136/amiajnl-2012-001073.

[20] E. Ammenwerth, W.O. Hackl, D. Riedmann, M. Jung, Contextualization of automatic alerts during electronic prescription: researchers' and users' opinions on useful context factors, Stud Health Technol Inform 169 (2011) 920–924, http://dx.doi.org/10.3233/978-1-60750-806-9-920.

[21] D.W. Bates, G.J. Kuperman, S. Wang, T. Gandhi, A. Kittler, L. Volk, et al., Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality, J Am Med Informatics Assoc 10 (2003) 523–530, http://dx.doi.org/10.1197/jamia.M1370.

[22] D.S. McEvoy, D.F. Sittig, T.-T. Hickman, S. Aaron, A. Ai, M. Amato, et al., Variation in high-priority drug-drug interaction alerts across institutions and electronic health records, J Am Med Informatics Assoc (2016).

[23] Zucchero F, Hogan M, Sommer C. Evaluations of drug interactions. St. Louis, Missouri, USA; 2005.

[24] A.L. Osornio, D. Luna, M.L. Gambarte, A. Gomez, G. Reynoso, F.G.B. de Quirós, Creation of a local interface terminology to SNOMED CT, Stud Health Technol Inform 129 (2007) 765–769.

[25] D. Luna, V. Otero, D. Canosa, S. Montenegro, P. Otero, F. González Bernaldo de Quirós, Analysis and redesign of a knowledge database for a drug-drug interactions alert system, CIN 129 (2007) 885–889.

[26] Luna D, Baum AJ, Giunti G, Tarsetti E, Díaz S, Briatore A. Coincidencias de pares de interacción droga-droga en trabajos seleccionados y el Sistema Notificador De Interacciones Farmacológicas del Hospital Italiano de Buenos Aires. Congr. Argentino Med. Interna, SAM. Sociedad Argentina de Medicina; 2013.

[27] R.T. Scheife, L.E. Hines, R.D. Boyce, S.P. Chung, J.D. Momper, C.D. Sommer, et al., Consensus recommendations for systematic evaluation of drug-drug interaction evidence for clinical decision support, Drug Saf 38 (2015) 197–206, http://dx.doi.org/10.1007/s40264-014-0262-8.

[28] H.M. Seidling, S. Phansalkar, D.L. Seger, M.D. Paterno, S. Shaykevich, W.E. Haefeli, et al., Factors influencing alert acceptance: a novel approach for predicting the success of clinical decision support, J Am Med Inf Assoc 18 (2011) 479–484, http://dx.doi.org/10.1136/amiajnl-2010-000039.

[29] D. Luna, M. Quispe, Z. Gonzalez, A. Alemrares, M. Risk, M. Garcia Aurelio, et al., User-centered design to develop clinical applications. Literature review, Stud. Health Technol. Inform. 216 (2015) 967, http://dx.doi.org/10.3233/978-1-61499-564-7-967.

[30] A.L. Russ, S. Chen, B.L. Melton, E.G. Johnson, J.R. Spina, M. Weiner, et al., A novel design for drug-drug interaction alerts improves prescribing efficiency, J Comm J Qual Patient Saf 41 (2015) 396–405.

[31] A. Sears, J.A. Jacko, Human-computer interaction: development process, CRC Press, 2009.

[32] V.L. Patel, T.G. Kannampallil, Cognitive informatics in biomedicine and healthcare, J Biomed Inform 53 (2015) 3–14, http://dx.doi.org/10.1016/j.jbi.2014.12.007.

[33] D.A. Norman, S.W. Draper, User centered system design: new perspectives on human-computer interaction, 1st ed., CRC Press, Hillsdale, N.J., 1986.

[34] ISO 92410-210:2010: Part 210: Ergonomics of human–system interaction: Human-centred design for interactive systems. ISO; 2010.

[35] A. Kushniruk, C. Nøhr, Participatory design, user involvement and health IT evaluation, Stud Health Technol Inform 222 (2016) 139–151.

[36] C. Macaulay, D. Sloan, X. Jiang, P. Forbes, S. Loynton, J.R. Swedlow, et al., Usability and user-centered design in scientific software development, IEEE Softw 26 (2009) 96–102, http://dx.doi.org/10.1109/MS.2009.27.

[37] B. Kitchenham, D.I.K. Sjoberg, T. Dyba, O.P. Brereton, D. Budgen, M. Host, et al., Trends in the quality of human-centric software engineering experiments – a quasi-experiment, IEEE Trans Softw Eng 39 (2013) 1002–1017, http://dx.doi.org/10.1109/TSE.2012.76.

[38] Kitchenham B, Budgen D, Brereton P, Turner M. 2nd International workshop on realising evidence-based software engineering (REBSE-2): Overview and Introduction. Second Int. Work. Realis. Evidence-Based Softw. Eng. (REBSE '07), IEEE; 2007, p. 1–1. doi:http://dx.doi.org/10.1109/REBSE.2007.1.

[39] D.I.K. Sjoeberg, J.E. Hannay, O. Hansen, V.B. Kampenes, A. Karahasanovic, N.-K. Liborg, et al., A survey of controlled experiments in software engineering, IEEE Trans Softw Eng 31 (2005) 733–753, http://dx.doi.org/10.1109/TSE.2005.97.

[40] R. Marcilly, L. Peute, M.C. Beuscart-Zephir, From usability engineering to evidence-based usability in health IT, Evidence-Based Heal Inform Promot Saf Effic Sci Meth Ethical Policy 222 (2016) 126–138, http://dx.doi.org/10.3233/978-1-61499-635-4-126.

[41] First hospital in Argentina awarded for digital maturity n.d. <https://www.eiseverywhere.com/file_uploads/b4ead7f8dcbcbe30ebebf601e62dfeaa_Argentina_HospitalItalianoPressRelease.pdf> [accessed September 27, 2016].

[42] Noticias Departamento de Informática en Salud del Hospital Italiano. Único Hospital en Argentina en acreditar calidad de Historia Clínica Electrónica 2016. <https://www1.hospitalitaliano.org.ar/infomed/index.php?contenido=ver_curso.php&id_curso=19752> [accessed September 27, 2016].

[43] Stanziola E, Uznayo MQ, Ortiz JM, Simón M, Otero C, Campos F, et al. User-centered design of health care software development: towards a cultural change. Stud Health Technol Inform, vol. 216. IOS Press; 2015. p. 368–71.

[44] Morae n.d. <https://www.techsmith.com/morae.html>.

[45] J. Simonsen, T. Robertson, Routledge international handbook of participatory design, Routledge, 2013.

[46] D. Luna, C. Otero, A. Almerares, E. Stanziola, M. Risk, F. Gonzalez Bernaldo de Quiros, Participatory design for drug-drug interaction alerts, Stud Health Technol Inform 210 (2015) 45–49, http://dx.doi.org/10.3233/978-1-61499-512-8-45.

[47] N.E. Schoenberg, H. Ravdal, Using vignettes in awareness and attitudinal research, Int J Soc Res Methodol 3 (2000) 63–74, http://dx.doi.org/10.1080/136455700294932.

[48] S.C. Evans, M.C. Roberts, J.W. Keeley, J.B. Blossom, C.M. Amaro, A.M. Garcia, et al., Vignette methodologies for studying clinicians' decision-making: validity, utility, and application in ICD-11 field studies, Int J Clin Heal Psychol 15 (2015) 160–170, http://dx.doi.org/10.1016/j.ijchp.2014.12.001.

[49] Soegaard M, Dam RF. The encyclopedia of human-computer interaction. Encycl Human-Computer Interact; 2012. <file:///books/hci.html>.

[50] Balsamiq n.d. <https://balsamiq.com/>.

[51] M.W.M. Jaspers, T. Steen, Bos C Van Den, M. Geenen, The think aloud method: a guide to user interface design, Int J Med Inform 73 (2004) 781–795, http://dx.doi.org/10.1016/j.ijmedinf.2004.08.003.

[52] Axure n.d. <http://www.axure.com/>.

[53] P.W. Jordan, B. Thomas, B.A. Weerdmeester, A.L. McClelland, SUS: a "quick and dirty" usability scale, Taylor and Francis, London, 1986.

[54] J. Brooke, SUS - a quick and dirty usability scale, Usability Eval Ind 189 (1996) 4–7, http://dx.doi.org/10.1002/hbm.20701.

[55] K. Zheng, K. Fear, B.W. Chaffee, C.R. Zimmerman, E.M. Karls, J.D. Gatwood, et al., Development and validation of a survey instrument for assessing prescribers' perception of computerized drug-drug interaction alerts, J Am Med Informatics Assoc 18 (2011) i51–i61, http://dx.doi.org/10.1136/amiajnl-2010-000053.

[56] R: The R Project for Statistical Computing n.d. <https://www.r-project.org/>.

[57] R. Marcilly, E. Ammenwerth, F. Vasseur, E. Roehrer, M.C. Beuscart-Zéphir, Usability flaws of medication-related alerting functions: a systematic qualitative review, J Biomed Inform 55 (2015) 260–271, http://dx.doi.org/10.1016/j.jbi.2015.03.006.

[58] A.B. McCoy, E.J. Thomas, M. Krousel-Wood, D.F. Sittig, Clinical decision support alert appropriateness: a review and proposal for improvement, Ochsner J 14 (2014) 195–202.

[59] S. Phansalkar, M. Zachariah, H.M. Seidling, C. Mendes, L. Volk, D.W. Bates, Evaluation of medication alerts in electronic health records for compliance with human factors principles, J Am Med Informatics Assoc 21 (2014) e332–e340, http://dx.doi.org/10.1136/amiajnl-2013-002279.

[60] A.L. Russ, A.J. Zillich, B.L. Melton, S.a Russell, S. Chen, J.R. Spina, et al., Applying human factors principles to alert design increases efficiency and reduces prescribing errors in a scenario-based simulation, J Am Med Inform Assoc (2014) 1–10, http://dx.doi.org/10.1136/amiajnl-2013-002045.

[61] H.M. Seidling, U. Klein, M. Schaier, D. Czock, D. Theile, M.G. Pruszydlo, et al., What, if all alerts were specific – estimating the potential impact on drug interaction alert burden, Int J Med Inform 83 (2014) 285–291, http://dx.doi.org/10.1016/j.ijmedinf.2013.12.006.

[62] A.T. Rodrigues, R. Stahlschmidt, S. Granja, A.L.E. Falcão, P. Moriel, P.G. Mazzola, Clinical relevancy and risks of potential drug–drug interactions in intensive therapy, Saudi Pharm J 23 (2015) 366–370, http://dx.doi.org/10.1016/j.jsps.2014.11.014.

[63] P.L. Smithburger, S.L. Kane-Gill, A.L. Seybert, Drug-drug interactions in the medical intensive care unit: an assessment of frequency, severity and the medications involved, Int J Pharm Pract 20 (2012) 402–408, http://dx.doi.org/10.1111/j.2042-7174.2012.00221.x.

[64] M. Ismail, F. Khan, S. Noor, I. Haider, I.-U. Haq, Z. Ali, et al., Potential drug-drug interactions in medical intensive care unit of a tertiary care hospital in Pakistan, Int J Clin Pharm 38 (2016) 1052–1056, http://dx.doi.org/10.1007/s11096-016-0340-3.

[65] E.K. Ahn, H.J. Kam, D.K. Park, E.Y. Jung, Y. Lee, R.W. Park, Differences among admitting departments in alerts and alert overrides for drug-drug interaction, Pharmacoepidemiol Drug Saf 23 (2014) 390–397, http://dx.doi.org/10.1002/pds.3591.

[66] V. Bhatt-Mehta, "Potential" drug-drug interactions and the picu: should we worry about ICU polypharmacy?*, Pediatr Crit Care Med 17 (2016) 470, http://dx.doi.org/10.1097/PCC.0000000000000694.

[67] A.W. Kushniruk, V.L. Patel, Cognitive and usability engineering methods for the evaluation of clinical information systems, J Biomed Inform 37 (2004) 56–76, http://dx.doi.org/10.1016/j.jbi.2004.01.003.

[68] L.A. Green, D. Nease, M.S. Klinkman, Clinical reminders designed and implemented using cognitive and organizational science principles decrease reminder fatigue, J Am Board Fam Med 28 (2015) 351–359, http://dx.doi.org/10.3122/jabfm.2015.03.140243.