Building an infrastructure to improve cardiac rehabilitation: from guidelines to audit and feedback

Verheul, M.M.

Publication date
2016

Document Version
Final published version

Citation for published version (APA):
OPTIMIZING THE USER INTERFACE OF A DATA ENTRY MODULE FOR AN ELECTRONIC PATIENT RECORD FOR CARDIAC REHABILITATION: A MIXED METHOD USABILITY APPROACH

Mariëtte M. van Engen-Verheul¹, Linda W.P. Peute¹, Nicolette F. de Keizer¹, Niels Peek¹,², Monique M.W. Jaspers¹

¹ Department of Medical Informatics, Academic Medical Centre/ University of Amsterdam, Amsterdam, The Netherlands
² MRC Health e-Research Centre, Institute of Population Health, University of Manchester, Manchester, UK

ABSTRACT

Introduction: Cumbersome electronic patient record (EPR) interfaces may complicate data-entry in clinical practice. Completeness of data entered in the EPR determines, among other things, the value of computerized clinical decision support (CCDS). Quantitative usability evaluations can provide insight into mismatches between the system design model of data entry and users’ data entry behaviour, but not into the underlying causes for these mismatches. Mixed method usability evaluation studies may provide these insights, and thus support generating redesign recommendations for improving an EPR system’s data entry interface.

Aim: To improve the usability of the data entry interface of an EPR system with CCDS in the field of cardiac rehabilitation (CR), and additionally, to assess the value of a mixed method usability approach in this context.

Methods: Seven CR professionals performed a think-aloud usability evaluation both before (beta-version) and after the redesign of the system. Observed usability problems from both evaluations were analyzed and categorized using Zhang et al.’s heuristic principles of good interface design. We combined the think-aloud usability evaluation of the system’s beta-version with the measurement of a new usability construct: users’ deviations in action sequence from the system’s predefined data entry order sequence. Recommendations for redesign were implemented. We assessed whether the redesign improved CR professionals’ 1) task efficacy (with respect to the completeness of data they collected), and 2) task efficiency (with respect to the average number of mouse clicks they needed to complete data entry subtasks).

Results: With the system’s beta version, 40% of health care professionals’ navigation actions through the system deviated from the predefined next system action. The causes for these deviations as revealed by the think-aloud method mostly concerned mismatches between the system design model for data entry action sequences and users expectations of these action sequences, based on their paper-based daily routines. This caused non completion of data entry tasks (31% of main tasks completed), and more navigation actions than minimally required (146% of the minimum required). In the redesigned system the data entry navigational structure was organized in a flexible way around an overview screen to better mimic users’ paper-based daily routines of collecting patient data. This redesign resulted in an increased number of completed main tasks (70%) and a decrease in navigation actions (133% of the minimum required). The think-aloud usability evaluation of the redesigned system showed that remaining problems concerned flexibility (e.g. lack of customization options) and consistency (mainly with layout and position of items on the screen).
**Conclusion:** The mixed method usability evaluation was supportive in revealing the magnitude and causes of mismatches between the system design model of data-entry with users’ data entry behaviour. However, as both task efficacy and efficiency were still not optimal with the redesigned EPR, we advise to perform a cognitive analysis on end users’ mental processes and behavior patterns in daily work processes specifically during the requirements analysis phase of development of interactive healthcare information systems.

**Key words:** Usability Evaluation; Electronic Health Records; CCDS; Cardiac Rehabilitation.
INTRODUCTION

The primary aim of recording data in electronic patient records (EPRs) is to support the delivery of good care, clinical decision-making, communication between healthcare workers and continuity of care. Additionally, EPRs are a valuable source of quality assurance of medical practice and scientific research [1]. In achieving these aims, effective use of EPRs requires structured data entry; which may be a challenge for physicians when design and implementation of an EPR does not align with their cognitive and workflow requirements and preferences [1-3]. Poorly designed and cumbersome data entry interfaces can complicate structured EPR data entry during clinical practice, resulting in poor data quality and data incompleteness [4, 5]. This may consequently lead to suboptimal functioning of health information technology systems integrated in the EPR, e.g. reminder systems, computerized physician order entry and computerized clinical decision support (CCDS).

Of those systems, CCDS is one of the most effective strategies to improve clinical decision making [4, 6]. CCDS uses characteristics of individual patients to generate patient-specific recommendations (based on national guidelines, evidence analysis or expert opinion) at the time and place clinical decisions are made [7]. To do so, CCDS systems often require availability of a large number of patient data (demographic data, data on complaints, symptoms, previous history, physical examination, laboratory, and other tests). Clinicians, health care staff, or patients can manually enter the data into the system; in addition, the EPR can be queried for retrieval of patient data [6]. Despite their goal to improve the quality of care, systematic reviews of CCDS studies reported only an improvement in professional performance for somewhat more than half of the included studies [8, 9] and attempts to identify critical success factors for CCDS systems have provided inconsistent results [8]. CCDS systems that derive their data form EPRs may provide inadequate advices as a result from incompleteness of EPR data needed to generate that advice [10].

Users of computerized systems are known to acquire knowledge about the system design models through experience that form the basis for the construction of reasonable action sequences. To stimulate complete data collection, an EPR systems’ design model of data entry (“the way the designer represents the system’s data entry functionality to the user, including screen presentations, interaction structure, and object relationships”) should match the users’ data entry behaviour (“the way that users have internalized how the data entry should proceed based on their experiences from daily practice”) [11]. Consequently, evaluation of the usability of the data entry interface in EPR systems is an essential step in human-centered design to optimize the match between the systems’ design model and user’s behaviour of data entry. Several quantitative methods exist (e.g. sequential pattern analysis, keystroke models and log file analysis) to analyse or model navigation patterns and action sequences from system users [12, 13]. These measures can provide insight into mismatches between the user’s behaviour and systems’ design model, but not into the underlying system design aspects causing these mismatches. Mixed method usability evaluation studies may provide this insight, resulting in
concrete redesign recommendations and finally in improved usability of a system’s data entry interface [14, 15].

An EPR system with CCDS functionalities, called MediScore CARDSS, was developed to stimulate guideline implementation on cardiac rehabilitation (CR) throughout the Netherlands [16, 17]. To guarantee complete data collection of the patient’s overall condition, a beta and a redesigned version of the system were both assessed by a mixed method usability evaluation with end-users (CR professionals). The results of the usability evaluation of the beta system version were handed over to the developers to improve the design of the system. The aim of this study was to improve the usability of the data entry interface of this first system version. Additionally we assessed the value of a mixed method usability approach (measuring fit between the systems’ design model of data entry and users’ data entry behaviour both from a quantitative and qualitative perspective) in this context.

BACKGROUND

Clinical setting: cardiac rehabilitation in the Netherlands
CR is a multidisciplinary therapy to support recovery from a cardiac incident or intervention, with the aim to improve a patient’s physical and psychological condition [18]. CR is recommended for all patients who have been hospitalized for an acute coronary syndrome (ACS) and for those who have undergone a cardiac intervention [19]. A meta-analysis shows consistent evidence of the effectiveness of exercise-based and multimodal (e.g., psychosocial and stress management) CR interventions with regard to mortality and prevention of future cardiac events (relative-risk reduction 21–47%) [20]. The therapy is offered by multidisciplinary teams, which generally include cardiologists, specialist nurses (of whom one acts as the rehabilitation coordinator), physical therapists, psychologists, dieticians and social workers, and is supported by a medical secretarial office.

Consistent with international guidelines, the Dutch guidelines for CR state that patients should be offered an individualized rehabilitation program based on their medical, physical, and psychosocial needs [21]. Traditionally this program is formulated during a 30 to 60 minute clinical patient interview, usually performed by a specialized nurse, physiotherapist or social worker. To structure the interview the guidelines include a paper-based clinical algorithm defining an extensive needs assessment procedure (NAP) [22]. This algorithm was designed in collaboration with CR professionals and is used in practice by multidisciplinary CR teams throughout the Netherlands [23]. It consists of fifteen numbered flowcharts across five domains, each describing how to select rehabilitation goals and therapies based on 155 to 175 patient data items (including both general questions and eight standardized questionnaires). During the daily routine with the paper-based NAP patient interview, professionals can adapt the order of data collection to their own preferences and as such data collection is flexible. A structured NAP to base therapy decisions on is a commonly used strategy within disease management of chronic patients [24]. It is needed to reduce inter-practice variation in the offered health care
and is in line with recommendations from the Chronic Care Model. This model is widely used to improve quality of care for chronic patients [24].

**Context: The MediScore CARDSS system**

The MediScore CARDSS system was developed in 2010 by ItéMedical BV, a Dutch commercial vendor in healthcare IT. The system concerns registration of administrative patient data, entry of clinical and health-related patient data, and provides CCDS to support CR professionals in the selection of goals and therapies for a guideline-based and patient-tailored CR program.

The design model of the beta system version implemented data entry in the exact order of the flowcharts of the Dutch paper-based clinical algorithm for CR. The entry of clinical and health-related data was partially static (e.g., standard questionnaires for quality of life and lifestyle assessments) and partially dynamic (i.e., the flow through the charts depends on previously entered data). To ensure complete data entry during the NAP, the beta system version guided users in one predefined data entry order through 53 data entry screens to collect all data items required to generate a patient-specific CR advice. Figure 1 shows one of the data entry screens of the beta system version. The flow through the system was supported with a ‘next button’ on each screen. Alternatively, navigation controls such as horizontal and vertical tabs displayed on both sides of the computer screen could be used when a user liked to deviate from the predefined data entry order. After entering all available patient data, the system provided its users with a patient-specific, guideline-based CR program, consisting of recommended rehabilitation goals and therapies. For each recommended goal and therapy professionals could either indicate that they adhere to the recommendation or that they did not follow the advice due to e.g. professional expertise, patient preferences, or lack of resources.
Figure 1 – Screenshot beta system version (MediScore CARDSS 2.0): NAP data entry concerning the patients’ social condition.

Figure 2 – Screenshot redesigned system version (MediScore CARDSS 3.0): grouped data entry of all static, standard questionnaires during the NAP.
CHAPTER 05

METHODS

Study design
We performed a think-aloud usability evaluation with CR professionals of the data entry interface of a beta (MediScore CARDSS 2.0) and the redesigned (MediScore CARDSS 3.0) version of the system. We combined the think-aloud usability evaluation of the system beta-version with the measurement of a new usability construct: users’ deviations in action sequence from the system’s predefined data entry order sequence. Results on users’ deviations from the systems’ design model of data entry order were combined with explanations for these deviations as revealed by the think-aloud method. These results and recommendations for redesign were handed over to the developers, who reengineered the beta version of the system.

We assessed whether the redesigned system version improved users’ task efficacy and task efficiency; with task efficacy defined as the number of data entry tasks and subtasks completed, and task efficiency defined as the average number of mouse clicks needed to do so. A data entry task was considered completed when a user succeeded in completion of each of its subtasks. As a measure of task efficiency, we calculated the difference between the theoretical minimum and actual number of mouse clicks users needed to complete each of the data entry subtasks successfully.

Participants
At the start of our research five CR centres (out of the 91 centres in the Netherlands of which 12 located in specialized rehabilitation clinics [25]) signed an intention with the software developers to buy the CARDSS Online system. They all received a demonstration of the beta system version and a local, stand-alone copy for explorative use. Two of them were located in non-teaching hospitals, one in a teaching hospital and two in rehabilitation clinics. All five centres were invited by mail to participate and to name one or two representatives of the multidisciplinary teams who were used to perform the paper-based NAP patient interview and who would work with the system in the future. The centres all agreed and by phone we invited the proposed seven participants to discuss details of the study (two centres proposed two participants). Characteristics of the participants and their centres are shown in Table 1. As the think-aloud method provides a rich source of data, a sample of approx. 8 subjects suffices to gain a thorough understanding of task behaviour or to identify the main usability problems with a computer system [26]). From July until September 2010 all seven participants performed the think-aloud of the beta system version. Results and identified usability problems of the beta system version were handed over to the software developers in December 2010 and presented elsewhere [27]. Reengineering by the developers took until March 2012. By that time two of the participating centres (one teaching and one rehabilitation clinic) had decided not to purchase the redesigned system and two of our participants were unavailable for the second think-aloud usability evaluation. To keep a convenient sample of subjects we invited two new participants two CR centres (one teaching and one non teaching hospital) who in meanwhile had procured
the redesigned system (but still did not actually use it in daily CR practice). From April until September 2012 five + two participants performed the think-aloud of the redesigned system version. Results and identified usability problems of the redesigned system version were handed over to the software developers in March 2013.

Table 1 – Characteristics of participants.

<table>
<thead>
<tr>
<th>Center ID</th>
<th>Center type</th>
<th>Discipline</th>
<th>Participation in think aloud</th>
<th>Beta system version</th>
<th>Redesigned system version</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Teaching</td>
<td>Nurse</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Non-teaching</td>
<td>Nurse</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Rehabilitation clinic</td>
<td>Social worker</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Rehabilitation clinic</td>
<td>CR medical secretary</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Non-teaching</td>
<td>Nurse</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Non-teaching</td>
<td>Social worker</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Teaching</td>
<td>Physiotherapist</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

Think-aloud user testing

We used the think-aloud method [28] to evaluate usability problems and analyse causes underlying users’ deviations in action sequence from the system’s predefined data entry order sequence. The think-aloud method is generally considered the ‘gold standard’ in usability evaluation as it provides detailed insight into user system interaction problems [29]. Seven CR professionals familiar with the paper-based patient documentation for the NAP, and with explorative system experience, performed a CR NAP in their own clinic by use of both the beta and redesigned system version. They were asked to enter data from 1) a fictitious patient case and 2) a real patient case from their own clinic (order of cases switched between participants to improve task completion for both cases). For the system usability evaluation with the fictitious patient case, users received all patient data in the predefined system data-entry order. With the real patient case, users were asked to perform the NAP with the system by entering data derived from a paper record from a patient recently treated in their own clinic. In both cases basic system functionalities were covered by asking the users to complete several main data entry tasks.

During the usability evaluation of the beta system version, users had to perform a complete CR NAP divided in seven main data entry tasks: patient registration, entering data concerning the patient’s physical condition, psychological condition, social condition, cardiovascular risk profile and lifestyle, and finally selecting goals and therapies for a patient-tailored CR program. Due to changes in the redesigned system version, the second evaluation concerned users’ to perform the same NAP, but now divided across five main tasks: patient registration, entering data on static, standard questionnaires and results of physical examination, entering dynamic...
data concerning the patients overall condition and finally the selecting of goals and therapies for the CR program. Table 4 and 5 give an overview of all main tasks defined for both system versions’ usability evaluations. Each main task itself was composed of several subtasks, e.g. defining the patient’s social condition required the entering of data about social functioning, the partner and work resumption. We identified a total of 41 subtasks per patient case for the beta system evaluation and 25 subtasks for the redesigned system evaluation. The amount of data items which needed to be entered during the think-aloud usability evaluation of both the beta and redesigned system version, was exactly the same. We performed 2-sample z-tests to compare sample proportions to assess whether task completion rates were different for the real and fictitious patient case and whether they improved in the redesigned system compared to its beta version.

**Assessment of navigation patterns and usability problems**

We used a mobile ‘usability lab’ consisting of a laptop with Morae™ software to capture screen, mouse gestures, keystrokes and the participant’s facial expressions and verbal reactions. Participants first performed a practice task to get accustomed with talking aloud before performing the two patient cases. All recorded data were analyzed with the Morae™ software. To assess the fit between the system’s design model with users’ data entry behaviour quantitatively in the beta system version, we measured the frequency with which users deviated from the system’s predefined data entry action through the CR NAP. As following the predefined data entry order sequence was important to attain data completeness, we used the number of deviations from the predefined system action sequence as a proxy of system quality on usability.

Data entry actions were first plotted in a graph to visualize individual navigational patterns for a thorough insight in the mismatches between the system’s design model of data entry and users’ data entry behavior. Thereafter, per user, a deviation percentage was calculated by dividing the number of actions in which he or she deviated from the next predefined system action by the total number of actions through the system. Herein, each of the 53 screens was considered one separate action. For instance, when a participant navigated through 50 screens in total, of which in 30 cases he or she went to the predefined next screen directly and in 20 cases to another screen, the deviation percentage was 40% (20/50). However, due to a completely flexible data entry structure in the redesigned system version, the construct used to assess navigational patterns could not be used again in the second usability evaluation.

To analyze the causes of participants’ deviations from the systems’ predefined data entry order qualitatively, all verbal protocols of participants were transcribed and verbal utterances and related video analysis were coded semi-bottom up by two researchers independently to reveal usability problems. Results of the researchers were compared to calculate inter-rater agreement (Cohen’s kappa statistic) which was interpreted according to the criteria published by Landis and Koch [30]. Thereafter discrepancies were resolved by discussion and, if needed, by consultation of a third researcher. Main usability problems revealed per participant were first listed by the two researchers independently to remove within and between participant
_duplicates of usability problems revealed. The usability problems, which both researchers rated as actual usability problems, were summed up into one master usability problem list. Two researchers subsequently classified each usability problem description in the master usability problem list according to the heuristic classification described by Zhang et al as summarized in Table 3 [31] and on severity according to the severity rating scale (1 – cosmetic problem only, 2 – minor usability problem, 3 – major usability problem, and 4 – usability catastrophe) defined by Nielsen et al [32]. The software developers also rated each problem on severity. Their rating was combined with the rating of the research team in a median rating (including the inter quartile range [IQR]). The team of researchers and software developers discussed these ratings of usability problems to determine priorities for the redesign of the system. To assess whether severity ratings for each individual heuristic class and the overall ratings over heuristic classes significantly differed for the beta vs. the redesigned system version, we performed a Mann Whitney U test. Per heuristic, we determined the number of usability problems in both the beta and redesigned system version. As recommended, we only performed the Mann-Whitney test for those heuristics with a sample size of more than seven problems [33].

RESULTS

Participants
All nine CR professionals had over three years general computer experience, seven of them were female, and four of them had used a previous version of the system (from 2004 [34]). Usability evaluations were performed within two weeks after participants had received the local, stand-alone copy of the system for explorative use. Participants involved in the usability evaluation were: four nurses, two social workers, a supportive CR medical secretary and two physiotherapists. Mean age of the professionals was 41.2 years (range 27 – 58). Overall these participants represented both the different types of CR centres (located in teaching and non-teaching hospitals and specialized rehabilitation clinics) and the multiple disciplines who are involved in the paper-based NAP patient interview.

Usability evaluation beta system version
Navigational patterns - On average, in the beta system version users deviated in 40% of the actions taken from the predefined next system action. Table 2 and Figure 3 give an overview of users’ deviations in action sequence from the system’s predefined data entry order sequence, for all 14 patient cases. Overall, users more often seemed to deviate from the predefined data entry order sequence during the data entry process of the NAP for the real patient case using the interview report from one of their own paper records (46%), than during the data-entry process for the fictitious patient case (36%). However, due to the high inter-individual variation shown in Figure 3, we were not able to derive one typical user model of deviations in action sequence from the system’s predefined data entry order sequence. Each participant followed his or her own specific route through the data entry process with only few process actions overlapping.
Table 2 – Number of users’ deviations in action sequence from the system’s predefined data entry order sequence, compared to the total number of actions users performed in the beta version of the system.

<table>
<thead>
<tr>
<th>User</th>
<th>Deviations/Total actions</th>
<th>%</th>
<th>Deviations/Total actions</th>
<th>%</th>
<th>Deviations/Total actions</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fictitious patient case</td>
<td></td>
<td>Real patient case</td>
<td></td>
<td>Overall</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>19 / 40</td>
<td>47.5</td>
<td>13 / 33</td>
<td>39.4</td>
<td>32 / 73</td>
<td>43.8</td>
</tr>
<tr>
<td>3</td>
<td>25 / 58</td>
<td>43.1</td>
<td>47 / 72</td>
<td>65.3</td>
<td>72 / 130</td>
<td>55.4</td>
</tr>
<tr>
<td>7</td>
<td>18 / 61</td>
<td>29.5</td>
<td>18 / 46</td>
<td>39.1</td>
<td>36 / 107</td>
<td>33.6</td>
</tr>
<tr>
<td>Sub total</td>
<td>62/159</td>
<td>39.0</td>
<td>78/151</td>
<td>51.7</td>
<td>140/310</td>
<td>45.2</td>
</tr>
<tr>
<td>2</td>
<td>29 / 46</td>
<td>63.0</td>
<td>27 / 47</td>
<td>57.4</td>
<td>56 / 93</td>
<td>60.2</td>
</tr>
<tr>
<td>4</td>
<td>14 / 61</td>
<td>23.0</td>
<td>26 / 56</td>
<td>46.4</td>
<td>40 / 117</td>
<td>34.2</td>
</tr>
<tr>
<td>5</td>
<td>25 / 86</td>
<td>29.1</td>
<td>7 / 53</td>
<td>13.2</td>
<td>32 / 139</td>
<td>23.0</td>
</tr>
<tr>
<td>6</td>
<td>14 / 47</td>
<td>29.8</td>
<td>14 / 27</td>
<td>51.9</td>
<td>28 / 74</td>
<td>37.8</td>
</tr>
<tr>
<td>Sub total</td>
<td>82/240</td>
<td>34.2</td>
<td>74/183</td>
<td>40.4</td>
<td>156/423</td>
<td>36.9</td>
</tr>
<tr>
<td>TOTAL</td>
<td>144 / 399</td>
<td>36.1</td>
<td>152 / 334</td>
<td>45.5</td>
<td>296 / 733</td>
<td>40.4</td>
</tr>
</tbody>
</table>

Main usability problems - Coding of the verbal think-aloud protocols of the beta system version resulted in a substantial agreement (kappa statistic of 0.61 [30]). Overall they revealed 45 main usability problems with a mean severity rate of 2.3. Table 3 gives an overview of the number of violations and a mean severity rate per heuristic. The heuristics most often violated were Consistency and standards (n=13) and Match between system and world (n=12). The heuristics with the highest severity rate (and violated more than once) were Users in control (3.5), Visibility of system state (3.0) and Match between system and world (2.9). An example of a usability problem classified as a Consistency and standards problem was “I do not know whether I need to use the button which says ‘Next’ or the button which says ‘Continue’ ”. An example of a Visibility of system state problem was “I navigate back to the previous page to check whether the data I entered was actually stored”.

A severe and frequently mentioned problem, that was visible in users’ deviations from the predefined data entry order, was that users were searching for a more flexible data entry order more in line with their daily working practices during the paper-based NAP. They stated they were searching for a flexible, but grouped data entry option to first enter all static, standard questionnaires before they would enter the dynamic patient data (e.g. “Now I have filled in the questionnaire on psychological functioning I’m searching for the other questionnaires, however I cannot find them on tab”). This problem was classified as Match between system and world.
Figure 3 – Users’ deviations in action sequence from the system’s predefined data entry order sequence in the beta version of the system.
Table 3 – Usability heuristics as adopted from Zhang et al [31], the number of violations (N) and a median severity rate with inter-quartile range (IQR) per heuristic both before (beta-version) and after the redesign of the system. The Mann Whitney U test was performed to test group differences for the ratings both per heuristic (only if sample size ratings was > 7) and overall from the usability evaluation of the beta and the redesigned system version.

<table>
<thead>
<tr>
<th>Heuristic</th>
<th>Explanation</th>
<th>Usability problems</th>
<th>Mann Whitney U test p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Beta system version</td>
<td>Redesigned system version</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>Median severity rate (IQR)</td>
</tr>
<tr>
<td>1 Consistency and standards</td>
<td>Users should not have to wonder whether different words, situations, or actions mean the same thing. Standards and conventions in product design should be followed.</td>
<td>13</td>
<td>2.5 (1.0 – 3.0)</td>
</tr>
<tr>
<td>2 Visibility of system state</td>
<td>Users should be informed about what is going on with the system through appropriate feedback and display of information.</td>
<td>2</td>
<td>3.0 (2.3 – 3.8)</td>
</tr>
<tr>
<td>3 Match between system and world</td>
<td>The image of the system perceived by users should match the model the users have about the system.</td>
<td>12</td>
<td>3.0 (2.3 – 3.0)</td>
</tr>
<tr>
<td>4 Minimalist</td>
<td>Any extraneous information is a distraction and a slow-down.</td>
<td>1</td>
<td>1.5 (1.0 – 2.0)</td>
</tr>
<tr>
<td>5 Minimize memory load</td>
<td>Users should not be required to memorize a lot of information to carry out tasks. Memory load reduces users’ capacity to carry out the main tasks.</td>
<td>1</td>
<td>3.0 (3.0 – 3.0)</td>
</tr>
<tr>
<td>6 Informative feedback</td>
<td>Users should be given prompt and informative feedback about their actions.</td>
<td>3</td>
<td>2.5 (1.0 – 3.0)</td>
</tr>
<tr>
<td>7 Flexibility and efficiency</td>
<td>Users always learn and users are always different. Give users the flexibility of creating customization and shortcuts to accelerate their performance.</td>
<td>3</td>
<td>1.5 (1.0 – 2.0)</td>
</tr>
<tr>
<td>8 Good error messages</td>
<td>The messages should be informative enough such that users can understand the nature of errors, learn from errors, and recover from errors.</td>
<td>3</td>
<td>2.0 (1.8 – 3.0)</td>
</tr>
<tr>
<td>9 Prevent errors</td>
<td>It is always better to design interfaces that prevent errors from happening in the first place.</td>
<td>1</td>
<td>3.0 (3.0 – 3.0)</td>
</tr>
<tr>
<td>10 Clear closure</td>
<td>Every task has a beginning and an end. Users should be clearly notified about the completion of a task.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11 Reversible actions</td>
<td>Users should be allowed to recover from errors. Reversible actions also encourage exploratory learning.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12 Use users’ language</td>
<td>The language should be always presented in a form understandable by the intended users.</td>
<td>2</td>
<td>1.0 (0.3 – 1.8)</td>
</tr>
<tr>
<td>13 Users in control</td>
<td>Do not give users that impression that they are controlled by the systems.</td>
<td>2</td>
<td>3.5 (3.0 – 4.0)</td>
</tr>
<tr>
<td>14 Help and documentation</td>
<td>Always provide help when needed, ideally context-sensitive help.</td>
<td>2</td>
<td>2.0 (2.0 – 2.0)</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>45</strong></td>
<td><strong>3.0 (2.0 – 3.0)</strong></td>
</tr>
</tbody>
</table>
Task efficacy - Table 4 shows that users on average completed 2.1 out of the 7 main tasks (30%) needed to enter data for a complete CR NAP successfully with the beta system version. Concerning the subtasks they completed 30 out of the 41 (73%) successfully. The subtasks concerning patient registration had the highest completion rate (86%). Subtasks with the lowest completion rates, concerned entering data referring to a patient’s cardiovascular risk profile (62%) and lifestyle (64%). A trend was found with users completing fewer tasks when they entered the data of a real patient case (19% main task, and 63% subtask completion) compared the situation where they had to enter data of the fictitious patient case (41% main task, and 82% subtask completion). However, in our sample the difference was neither statistically significant for the main task (proportion test 95% CI -0.26 – 0.70; p= 0.3691) nor for the subtask completion (proportion test 95% CI -0.28 – 0.66; p= 0.426).

Task efficiency - With the beta system version, users on average needed 321 mouse clicks to complete the subtasks for the NAP for one patient; that is 146% (range 108% - 245%) of the minimum number of mouse clicks (241). For the fictitious patient case this was 156% (range 125% - 245%) of the minimum number of mouse clicks, and for the real patient case 136% (range 108% - 179%).

Usability evaluation redesigned system version

Redesign - The redesigned system version basically provides the same functionalities as the beta system version, although the design of the data entry interface was significantly changed. Both individual navigation patterns and the think-aloud protocols from the beta system evaluation showed that the systems’ design model of one predefined data entry order did not match users’ expectations of, and their actual navigational patterns through the system. Users continually searched for, and expressed their preferences of ways to enter the data in the system in more flexible manners. Users indicated that in their daily routine with the paper-based NAP they tend to collect patient data in various manners, depending on their individual workflow. Users were thus unable to construct a mental model that matched the system’s design model of one predefined data entry order. The mixed method usability evaluation revealed a need for a completely flexible data entry navigational structure of the interface.

The redesigned system provides 32 data entry screens to guide the user through the CR algorithm. Users can decide on a grouped data entry of all static, standard questionnaires (although this is not required) before entering the dynamic data patient data and receiving the decision support on the final patient-specific rehabilitation program. Figure 2 shows the data entry screens for all standardized questionnaires in the redesigned system version. Complete data collection is stimulated by showing users which data entry actions they already have finished and which actions they still need to complete. The ‘next button’ on each screen was removed and the number of horizontal and vertical tabs reduced. In the redesigned system version, users are supported in their flow through the system by guiding them back to an overview screen after they have entered and saved data. From this overview screen, users can
choose the next data item or questionnaire they prefer to enter. As a result of this fully flexible
data entry interface user deviations in action sequence from the system’s predefined data entry
order sequence could not be assessed in the redesigned system.

**Main usability problems** - Coding of the verbal think-aloud protocols of the redesigned system
version resulted in a moderate agreement (kappa statistic of 0.53 [30]). The think-aloud usability
evaluation of the redesigned system revealed 30 usability problems with an overall median
severity of 2.0. This overall median severity was significant lower than the 3.0 overall median
severity rating in the beta system version (Mann Whitney U test p= 0.004). Table 3 gives an
overview of the number of violations and a mean severity rate per heuristic. The heuristics
most often violated were Consistency and standards (n=6), Flexibility and efficiency (n=6) and
Minimize memory load (n=5). The heuristics with the highest severity rate (and violated more
than once) were Minimize memory load (2.3), Prevent errors (2.3) and Use users’ language
(2.0). Problems related to incomplete data collection occurred in all heuristic classes e.g. in
Consistency and standards (“I know that I still miss some dynamic data on the patients’ physical
functioning, however I do not see where I can enter this data”), Flexibility and efficiency (“Why
do I need to enter the same date for each questionnaire again”) and Visibility of system state
(e.g. “I doubt whether the system is still saving data, as I have clicked on the ‘Save’ button but
I’m not automatically redirected to the overview page”). Three usability problems were classified
as violations of the heuristic Match between system and world with a mean severity of 1.5. A
recurrent theme over all heuristics was lack of customization options for e.g. system default
values (“We always plan the same amount of training sessions”) and terminology (“We use the
term acute coronary syndrome which I cannot find in the list”).

**Task efficacy** - Table 5 shows that users on average completed 3.8 out of the 5 main tasks
(76%) needed to perform a complete CR NAP successfully with the redesigned system version.
Compared to completion of main tasks in the beta system version (30%), the improvement
was statistically significant (proportion test 0.46; 95% CI 0.05 – 0.87; p= 0.0263). Concerning
the subtasks they completed 20 out of the 25 (79%) successfully. Compared to completion of
subtasks in the beta system version (73%), the improvement was not statistically significant
(proportion test 0.06; 95% CI -0.29 - 0.41; p= 0.7362). The subtasks concerning the selection
of goals and therapies for the patient specific CR program (93%) and patient registration (91%)
had the highest completion rates. The subtasks with the lowest completion rates were entering
dynamic data concerning the patients overall condition (55%) and results of a patient’s physical
examination (79%). A trend was found with users completing fewer subtasks when they
entered the data of a real patient case (31% completion for main tasks and 70% for subtasks)
compared to the situation where they had to enter data of the fictitious patient case (77%
main task and 87% subtask completion). However, in our sample this difference was neither
statistically significant for the main task (proportion test 0.46; 95% CI -0.16 – 1.08; p= 0.1445)
nor for the subtask completion (proportion test 0.17; 95% CI -0.34 – 0.68; p= 0.5129).
Task efficiency - With the redesigned system version, users on average needed 267 mouse clicks to complete subtasks for the NAP for one patient: 133% (range 106% - 216%) of the minimum number of mouse clicks (n=201). For the fictitious patient case this was 129% (range 117% - 193%) of the minimum number of mouse clicks, and for the real patient case 138% (range 106% - 216%).

Table 4 – Results of the usability evaluation of beta system version: Tasks and subtasks completed and mouse clicks needed.

<table>
<thead>
<tr>
<th>Beta system version</th>
<th>Task 1 Patient registration</th>
<th>Task 2 Data physical condition</th>
<th>Task 3 Data psychological condition</th>
<th>Task 4 Data social condition</th>
<th>Task 5 Data cardiovascular risk profile</th>
<th>Task 6 Data lifestyle</th>
<th>Task 7 Selection goals + therapies</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td># subtasks per task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># mouse clicks minimally required</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>4</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>59</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>54</td>
<td>339</td>
</tr>
</tbody>
</table>

Fictitious patient case

<table>
<thead>
<tr>
<th>Task completion</th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
<th>Task 5</th>
<th>Task 6</th>
<th>Task 7</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average subtask completion</td>
<td>96%</td>
<td>83%</td>
<td>98%</td>
<td>89%</td>
<td>73%</td>
<td>70%</td>
<td>75%</td>
<td>82%</td>
</tr>
<tr>
<td>Average mouse clicks needed for completed subtasks</td>
<td>151%</td>
<td>114%</td>
<td>109%</td>
<td>164%</td>
<td>240%</td>
<td>131%</td>
<td>182%</td>
<td>156%</td>
</tr>
</tbody>
</table>

Real patient case

<table>
<thead>
<tr>
<th>Task completion</th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
<th>Task 5</th>
<th>Task 6</th>
<th>Task 7</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average subtask completion</td>
<td>75%</td>
<td>62%</td>
<td>61%</td>
<td>71%</td>
<td>51%</td>
<td>59%</td>
<td>75%</td>
<td>63%</td>
</tr>
<tr>
<td>Average mouse clicks needed for completed subtasks</td>
<td>145%</td>
<td>108%</td>
<td>139%</td>
<td>131%</td>
<td>194%</td>
<td>117%</td>
<td>115%</td>
<td>136%</td>
</tr>
</tbody>
</table>
Table 5 – Results of the usability evaluation of redesigned system version: Tasks and subtasks completed and mouse clicks needed.

<table>
<thead>
<tr>
<th>Redesigned system version</th>
<th>Task 1: Patient registration</th>
<th>Task 2: Data psychosocial questionnaires</th>
<th>Task 3: Data physical examination</th>
<th>Task 4: Data overall condition</th>
<th>Task 5: Selection goals + therapies</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td># subtasks per task / # mouse clicks minimally required</td>
<td>3/8</td>
<td>4/37</td>
<td>6/21</td>
<td>4/105</td>
<td>25/303</td>
<td>30</td>
</tr>
<tr>
<td>Fictitious patient case</td>
<td>Task completion</td>
<td>6/7</td>
<td>7/7</td>
<td>7/7</td>
<td>1/7</td>
<td>6/7</td>
</tr>
<tr>
<td>Average subtask completion</td>
<td>95%</td>
<td>100%</td>
<td>100%</td>
<td>52%</td>
<td>96%</td>
<td>87%</td>
</tr>
<tr>
<td>Average mouse clicks needed for completed subtasks</td>
<td>120%</td>
<td>117%</td>
<td>128%</td>
<td>156%</td>
<td>193%</td>
<td>129%</td>
</tr>
<tr>
<td>Real patient case</td>
<td>Task completion</td>
<td>4/7</td>
<td>2/7</td>
<td>0/7</td>
<td>0/7</td>
<td>5/7</td>
</tr>
<tr>
<td>Average subtask completion</td>
<td>86%</td>
<td>70%</td>
<td>57%</td>
<td>57%</td>
<td>89%</td>
<td>70%</td>
</tr>
<tr>
<td>Average mouse clicks needed for completed subtasks</td>
<td>106%</td>
<td>108%</td>
<td>116%</td>
<td>216%</td>
<td>201%</td>
<td>138%</td>
</tr>
</tbody>
</table>

DISCUSSION

Statement of principal findings

We performed a mixed method usability evaluation of a data entry module of an EPR system with CCDS in the field of CR. We combined the think-aloud method with the new construct ‘users’ deviations in action sequence from the systems’ predefined data entry order sequence’. With the systems’ beta version, healthcare professionals deviated strongly and with high interindividual variation from the systems’ predefined data entry order. The think-aloud evaluation of the beta system version showed that users’ expectations of action sequences to be performed in the system, which were based on their daily paper-based patient documentation routines, mismatched the system design model of one predefined data entry order (mismatch between system and world). Other usability issues concerned e.g. consistency in button names, colors and position of data entry fields and visibility of saved data items. Given these mismatches between the beta version system design and the users’ daily paper-based documentation routines, they probably were hindered in the construction of an adequate mental model of the system (the way the user perceives that the system works based on his mental processes) [11]. Possibly, due to the fact that CR professionals did not construct such a mental model of the system, they may have been less capable to use the system properly and vice versa. Low system usability may have hindered users in the construction of an adequate mental model of the system. The mismatch resulted in incompletion of data-entry tasks and a higher number...
of navigation actions than required. In the redesigned system version, usability issues were solved by the designers within available resources. To better adapt to users’ expectations and support the construction of adequate mental models by users on how the system works, the data entry navigational structure was organized completely flexible around an overview screen. With the redesigned system, users completed more data-entry tasks while they performed less navigational actions. The think-aloud evaluation of the redesigned system revealed remaining usability problems. They mainly concerned lack of flexibility (e.g. customization options related to defaults and terminology) and still problems with consistency (mainly with layout and position of items on the screen).

**Strengths and weaknesses of the study**

A strength of our study is that we used a mixed method approach which combines the think-aloud method with the new usability construct on users’ deviations. Using this construct we were able to visualize that all users deviated strongly from the systems’ predefined data entry order, however all differently. This showed that although all users struggled to construct their mental model on data entry in the system, they did not share a common or general mental model on data entry. By combining measurements of this new usability construct with the analyses of the think-aloud user sessions, we revealed the usability problems causing users’ deviations. We obtained a detailed insight into the mismatches in the user’s data entry behaviour and the systems’ design model of data entry. This mixed approach allowed us to infer specific recommendations to optimize the systems’ design model for meeting users’ differing expectations concerning data entry order and supporting users in the construction of an adequate mental model of the system’s data entry modes.

Other strengths of our study are that we assessed CR professionals’ task efficacy and efficiency and number and severity of usability problems detected both before and after, so in the beta and redesigned system version. Hence, the results of this combination of quantitative and qualitative measurements revealed additional necessity and recommendations for redesign in the before study. Though our study (purposely) was more qualitative than quantitative in nature, our results indicate that the mixed method we applied to direct our redesign efforts was supportive in improving the data entry interface of the new system version.

Our study has some limitations as well. First, the tasks we defined for the usability evaluations were based on the guidelines for the CR NAP, which require the entry of an enormous set of patient data items (155-175). Entering such an amount of data asks for a strong commitment of the professional to finish this task. This challenges the design of the data entry interface compared to situations wherein users are supposed to enter a smaller amount of patient data. Second, the inter-rater agreement of the coding of verbal protocols of the redesigned system evaluation was only moderate (compared to a substantial inter-rater agreement of coding of verbal protocols concerning the beta system evaluation). More extensive pilot testing and discussion between the two researchers responsible for coding in the beginning of the coding process might have resulted in a higher inter-rater agreement. However, as explained, the
disagreements in coding were resolved by consultation with a third researcher. Together, causes underlying these differences in coding were revealed, after which final consensus on codes could easily be reached. Third, the Dutch clinical algorithm for the CR NAP on which the beta system version was based, was revised in 2010 and during the time of the usability evaluation not yet completely implemented in most Dutch CR clinics. The usability issues revealed may likewise be caused by system users’ unfamiliarity with the content of the algorithm. Fourth, usability issues concerning flexibility and consistency remained after the system’s redesign effort. Although we shared all our findings and recommendations for redesign with the system developers, we did not demonstrate our recommendations graphically through mock ups. A provision of mock-ups could possibly have assisted the developers in preventing possible misinterpretations of our redesign proposals, and usability issues remaining in the redesigned system. Finally, we did not evaluate the usability of the CARDSS system in daily practice, during clinical interviews of CR professionals with their patients. Field-based evaluation of the CARDSS system after its implementation in Dutch CR clinics may therefore reveal additional usability issues concerning mismatches of the system design model with CR professionals’ workflow. We aim to perform such an ‘in vivo’ evaluation of CARDSS as part of our future work focussing on improving CR care quality in the Netherlands. For such a field-based usability evaluation of CARDSS, we will combine observations of professionals in daily practice with screen logging, synchronous audio recording of CR professionals verbal utterances and semi-structured interviewing with the CR professionals.

Strengths and weaknesses in relation to other studies

The mixed method approach we used allowed for both a quantitative analysis of suboptimal navigational patterns of users through the system and a qualitative analysis of explanation of the usability issues causing these patterns. While we used the new construct ‘users’ deviations in action sequence from the system’s predefined data entry order sequence’, other usability studies of health information systems have applied somewhat similar quantitative methods to gain insight into health professionals’ systems’ navigational patterns in relation to their task performance [13, 35, 36]. Zheng et al [13] used sequential pattern analysis as a quantitative measure to uncover recurring user interface navigational patterns in an EPR system, whereas Lacerda et al. [35] used a keystroke-level model (KLM) to assess the efficiency of action sequences in two interfaces with differing interaction modes of a telecardiology system. Further, Kelders et al. [36] applied log file analysis in combination with a survey and real time usability tests to assess the usability of a web-based health intervention. A shortcoming of quantitative methods when applied without qualitative methods is however, that they do not provide detailed insight into the usability issues causing the mismatches between the users’ data entry behaviour and the system’s design model of data entry that led to suboptimal navigation patterns and task incompleteness.

To reveal usability problems resulting from the mismatch between predefined system action sequences and the daily working routines of healthcare professionals, several other
usability studies of health information systems have applied qualitative methods only [5, 37-39]. For instance Peute et al [37] performed a qualitative study using the cognitive walkthrough and think-aloud user sessions to reveal usability problems in a computerized physician ordering (CPOE) system. Second Beuscart et al [38] performed a heuristic evaluation and both in-lab and on-site usability tests for the evaluation of a CPOE medication system. Finally Rose et al [39] conducted multiple focus group and field study sessions with primary care physicians to improve the usability of a management module of a widely deployed web-based EPR. In all these, qualitative usability methods were useful in revealing usability problems concerning mismatches of the systems’ design model and the users’ system behaviour in various types of healthcare information systems. However, we combined our qualitative think-aloud method with the construct on users’ deviations from predefined system action sequences (see Figure 3). While the details on the usability problems as revealed by the think-aloud sessions were helpful in defining proposals for the system’s redesign, the visualization of the construct on users’ deviations was especially useful in clarifying the mismatches of users’ navigational patterns with the system design model of data entry to the system designers. In a single glance the figure pointed out the need for an update of the system data entry model.

Furthermore, although the studies using qualitative methods only, succeeded in proposing specific recommendations for a system redesign, none of them included a usability analysis afterwards. In our study we did perform a mixed-method usability analysis of the redesigned system version to gain insight in its usability and CR professionals’ performance in terms of their task efficiency and efficacy.

Meaning of the study
Our mixed method usability approach appeared helpful in improving the usability of the redesigned MediScore CARDSS system. The improved system’s design model better fitted each users’ daily documentation routines. As a consequence, CR professionals were better supported in building an adequate user model of how the redesigned system works. However, the think-aloud user sessions with the redesigned system version revealed that both task efficiency (completeness of data collection) and task efficacy (number of mouse clicks) were still not optimal as well as that usability problems remained. As we found recurrent usability problems in the redesigned system version related to e.g. consistency and flexibility, visualisation of proposed design solutions could have made our recommendations more clear. Concrete visual examples could have been a better way of communicating both usability problems and our recommendations to the system designers. Healthcare professionals seem to balance technological and work demands, all in an efficient and cost effective way. When confronted with usability problems, they may see a greater need to work around intended work practices and search for alternative routes in the system to perform their tasks (workarounds) [40, 41]. While we intended to make the beta system version more usable by making it more flexible, we may likewise have increased the options for CR professionals to work around certain data entry tasks. However, since the redesigned system version alerted the CR professionals on
uncompleted data entry tasks, such possible workarounds presumably did not negatively impact their final task efficiency and efficacy. Further, closely observing CR professionals in their daily routines of NAP data collection in the requirements analysis phase could have prevented some essential usability flaws in the beta system version. These observations would at least have revealed the variations in NAP data collection practices and the need for a flexible system data entry mode to accommodate CR professionals’ individual practices during paper-based NAP.

Using the think-aloud method during the requirements analysis phase, Jaspers et al [42] developed a cognitive task model reflecting pediatric oncologists’ task behavior in searching through a paper-based patient record as input for a prototype user interface of an EPR system. Likewise, Verhoeven et al [43] analyzed using the think-aloud method the processing of paper-based infection control guidelines by health care workers, to develop a user-oriented website for the communication of these guidelines. Finally Kilsdonk et al [44] developed an information processing model through an analysis of think-aloud protocols and used the model as input for the design of a CCDS user interface to support pediatric oncologists.

Based on these three studies, we advise to apply cognitive methods like the think-aloud method to analyze end users’ mental processes and behavior patterns in daily work processes specifically during the requirements analysis phase of development of interactive healthcare information systems such as EPRs. These insights may support designs of interactive healthcare information systems that are consistent with the users’ daily working routines [11, 42]. As a consequence, such system designs may simplify the construction of users’ mental models on how the system works.

CONCLUSION

The mixed method usability evaluation was supportive in revealing the magnitude and causes of mismatches between the system design model of data-entry with users’ data entry behaviour. However, as both task efficacy and efficiency were still not optimal with the redesigned EPR, we advise to perform a cognitive analysis on end users’ mental processes and behavior patterns in daily work processes specifically during the requirements analysis phase of development of interactive healthcare information systems such as EPRs. These insights may be of help in designing interactive healthcare information systems that map on users’ daily working routines, which may simplify the construction of users’ mental models on how the system works. Such an approach may ultimately lead to more efficient and effective healthcare information systems and more satisfied users.
SUMMARY TABLE

What was already known on the topic
- The value of computerized clinical decision support (CCDS) within an electronic patient record (EPR) is, among other things, determined by the completeness of data entered into the system.
- Mixed method usability evaluation studies may provide insight into mismatches between the system design model of data entry and users’ data entry behaviour to optimize the usability of a system’s data entry interface.

What this study added to our knowledge
- A mixed method usability evaluation combining the think-aloud method with the new usability construct ‘users’ deviations in action sequence from the system’s predefined data entry order sequence’ was helpful to visualize and name usability problems for user deviations from the systems’ predefined data entry order.
- The mixed method approach was supportive in revealing the magnitude and causes of mismatches between the system design model of data-entry with users’ data entry behaviour.
- CR professionals completed more main tasks with fewer navigation actions with the revised version of the system.

AUTHOR’S CONTRIBUTIONS

MvEV, NP and NdK had the basic idea for this study. Together with LP and MJ they were involved in the development of the protocol. MvEV performed the usability evaluations and processed the think-aloud protocols. MvEV, MJ and LP drafted the manuscript. All authors were involved in the critical revision of the paper for intellectual content. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The authors would like to thank the CR professionals, who invested their time to participate in our research, for their contribution. Furthermore we would like to thank the vendor of the MediScore CARDSS system, Ité Medical B.V., for collaboration. They made both a beta and a redesigned version of their system available for our mixed method usability evaluation with the CR professionals. Finally we thank Ellen Kilsdonk and Veerle Kilsdonk (both research assistants at the Department of Medical Informatics, Academic Medical Center, University of Amsterdam) for their assistance during the coding of the think-aloud transcriptions.
REFERENCES


37. Peute LW, Jaspers MW. The significance of a usability evaluation of an emerging laboratory order entry system.


