

Usability Evaluation of a Guideline Implementation System for Cardiac Rehabilitation: Think Aloud Study

M.M. van Engen-Verheul^a, L.W. Peute^a, E. Kilsdonk^a, N. Peek^a, M.W.M. Jaspers^a

^a Department of Medical Informatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

Abstract

Guidelines on cardiac rehabilitation state that a patient-tailored, comprehensive cardiac rehabilitation programme should be constructed for each patient based on a structured needs assessment procedure. We performed a usability evaluation with seven naive end-users of the MediScore CARDSS 2.0 system which implements such a procedure based on the Dutch guidelines. The analysis showed that users deviated strongly from the predefined data entry order; could not complete all subtasks for a complete needs assessment procedure, and needed far more navigation actions than minimally required. We conclude that the conceptual model of systems which implement guidelines requiring data entry should adapt to users' mental model concerning data entry order preferences to guarantee complete data collection.

1. Introduction

Cardiac rehabilitation (CR) is a multidisciplinary therapy to support heart patients recover from a cardiac incident or intervention, and aims to improve their overall physical and mental functioning [1]. Consistent with international guidelines, the Dutch guidelines for CR state that patients should be offered an individualized rehabilitation programme based on their medical, physical, and psychosocial needs [2]. To this end the guidelines are supplemented with a paper-based clinical algorithm that describes a structured needs assessment procedure (NAP) [3]. This algorithm was designed in collaboration with CR professionals and is used in practice by rehabilitation nurses and physiotherapists [4]. It consists of fifteen numbered flowcharts across five domains, each describing how to select rehabilitation goals and therapies based on collected patient data. During paper-based NAP the order of data collection is flexible and can be adapted to professionals' own preferences. According to the guidelines complete data collection is important though to construct a patient-tailored, comprehensive rehabilitation programme. Patients should only receive rehabilitation therapies that they really need, and not others. A structured NAP should also reduce interpractice 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 [5].

To stimulate implementation of the CR guidelines in the Netherlands, an electronic patient record system with computerized decision support facilities, called MediScore CARDSS 2.0, was recently developed by ItéMedicel BV, a Dutch commercial vendor in healthcare IT. The system has evolved from an earlier system, developed in 2004 by the University of Amsterdam, that was based on similar but less extensive guidelines. In a cluster randomised trial, it was shown that the previous system increased adherence of healthcare professionals' decisions with guideline recommendations [6].

To assess the new system's design for performing the CR NAP on usability, a beta version of MediScore CARDSS 2.0 was made available for an evaluation with naive end-users. The aim of this study was to evaluate 1) task efficacy of the system with respect to completeness of data collection; 2) fit between the system's model of predefined data entry order with users' mental models, and 3) task efficiency of the system with respect to performing all subtasks needed to complete a CR NAP.

2. System background

The MediScore CARDSS 2.0 system is based upon the Dutch clinical algorithm for CR. A complete data collection during the NAP is supported by guiding users in 53 screens through all the subtasks of the algorithm. The subtasks concern registration of administrative patient data, entry of clinical and health-related patient data, and finally the selection of goals and therapies for a patient-tailored rehabilitation programme. The entry of clinical and health-related data is partially static (e.g., standard questionnaires for quality of life and lifestyle assessments) and partially dynamic (i.e., the flow through data items depends on previously entered data). Figure 1 shows one of the data entry screens of the system.

The order of the data entry subtasks in the system follows the order of the algorithm's paper flowcharts and can be pursued with a 'next button' on each screen. Alternatively, navigation controls such as the vertical tabs displayed on the both sides of the computer screen can be used to deviate from the predefined data entry order. After entering all available data the system provides its users with a patient-specific, guideline-based rehabilitation programme, consisting of recommended rehabilitation goals and therapies. For each goal and therapy users can either select that they adhere to the recommendation or that they deviate because of e.g. professional expertise, patient preferences, or lack of resources. Furthermore the programme incorporates an electronic patient record for CR, background information from the guidelines and some additional functions specific to the working procedures in the multidisciplinary outpatient care.

3. Methods

We used the think aloud method to evaluate problems end-users encountered in interaction with the system [7]. Seven representative, naive end-users performed a CR NAP on their own workstation by entering data from 1) a fictitious patient case and 2) a real patient case from their own clinic. In both cases basic system functionalities were covered by asking the users to complete seven main 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 rehabilitation programme. Each main task is itself composed of several subtasks, e.g. the patient's social condition consists of entering data about social functioning, the partner and work resumption. We identified a total of 41 subtasks. For the system usability evaluation with the fictitious patient case, users received all patient data in the predefined system order. In the real patient case, users were asked to perform the NAP by entering data derived from a paper record from a patient recently treated in their own clinic in the system.

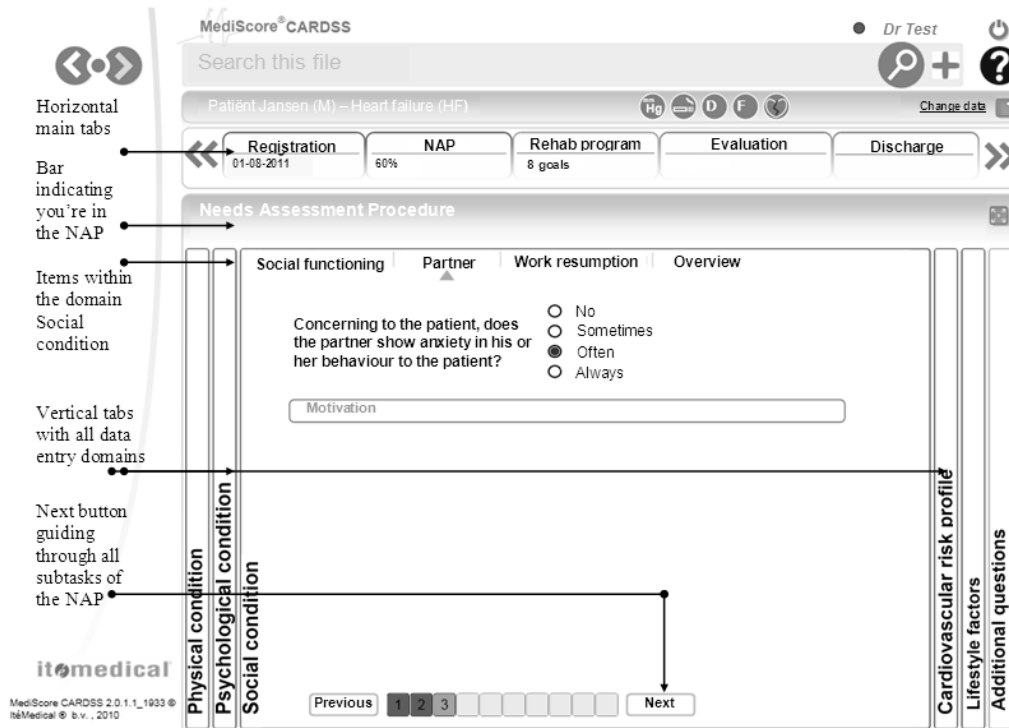


Fig. 1. Screenshot MediScore CARDSS 2.0: Data entry during the NAP concerning the social condition of the patient.

We used a mobile ‘usability lab’ consisting of a laptop with MoraeTM 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 starting with the two cases. All recorded data were analyzed with the MoraeTM software. We assessed the: 1) the number of tasks and subtasks successfully completed by each user; 2) the frequency with which users deviated from the system’s predefined route through the CR NAP, and 3) the difference between the theoretical minimum and actual number of mouse clicks users needed to complete each of the subtasks successfully. A task was considered completed when each of its subtasks was completed; partial completion of tasks was not possible. Associations between the number of mouse clicks, deviation from the predefined data entry order, and task completion was investigated by linear regression analysis.

4. Results

4.1 Number of tasks successfully completed

Table 1 gives an overview of the characteristics of the seven end users. All users had more than three year of general computer experience, most of them were female nurses, and about half of them had used the previous version of the system. Table 2 shows that on average users successfully completed 2.2 out of the 7 main tasks (30%) needed to perform a complete CR NAP. 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 were entering patient data concerning the cardiovascular risk profile (62%) and lifestyle (64%). Fewer subtasks were completed when users entered the data of a real patient case (63%) compared to data entry of the fictitious patient case (82%).

Table 1. User characteristics

| User | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|-------|-------|-------|---------------|-----------|-------|------------------|
| Sex (Female/ Male) | F | F | F | F | M | F | M |
| Age (years) | 27 | 41 | 48 | 23 | 38 | 58 | 55 |
| Discipline | Nurse | Nurse | Nurse | Social worker | Secretary | Nurse | Physio-therapist |
| Previously worked with CARDSS 1.0 | Yes | Yes | Yes | No | No | Yes | No |
| # patients previously entered in MediScore CARDSS 2.0 | 5-15 | >25 | 1-5 | 5-15 | 1-5 | 0 | 1-5 |
| General computer experiences (years) | > 3 | > 3 | > 3 | > 3 | > 3 | > 3 | > 3 |

4.2 Deviations from the predefined order of the NAP in the system

On average, users deviated in 41% of the steps taken from the predefined next system step. The deviation was larger during the data entry process of the NAP for the real patient case (45%), than in the data-entry process for the fictitious patient case (38%). Regression analysis showed that users who deviated from the predefined data entry order also completed fewer tasks. On average, each 40 deviations were associated with one task less being completed.

4.3 Differences between the theoretical minimum and actual number of mouse clicks

For completed subtasks for the NAP for one patient, users needed on average 321 mouse clicks: 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, and for the real patient case 136% (range 108% - 179%). Regression analysis showed that users with the highest number of mouse clicks had higher deviation rates from the predefined order of system tasks than users who followed the predefined order. On average, each deviation was associated with six mouse clicks.

5. Discussion and Conclusion

Computer-based guideline implementation systems that provide patient-specific recommendations based on data entry can improve guideline adherence and patient outcomes [8]. To stimulate optimal implementation of these systems and the underlying guidelines, the systems' conceptual models should match the mental model of their users [9]. In this study we found that the MediScore CARDSS 2.0 system's model of predefined order of system data entry tasks for the CR patient NAP does not fit the mental model of users as they deviated strongly from this order. They could not complete all data entry tasks as defined in the CR guidelines and needed far more navigation actions than was minimally required to perform a complete NAP. Particularly when users entered data concerning a real patient case, they followed a different data entry ordering sequence than predefined by the system.

By asking representative end users to perform a complete patients' NAP in their own work environment, the study provided insights into how MediScore CARDSS 2.0 design can be improved to better follow the task strategies of its users. The construct 'deviation from the

Table 2. Results of the usability evaluation: Tasks and subtasks completed and mouse clicks needed for completed subtasks

| | Task 1: Patient registration | | Task 2: Data entry physical condition | | Task 3: Data entry psychological condition | | Task 4: Data entry social condition | | Task 5: Data entry cardiovascular risk profile | | Task 6: Data entry lifestyle factors | | Task 7: Selecting rehabilitation programme | | TOTAL | |
|---|------------------------------------|--------|--|--------|---|--------|--|--------|---|--------|---|--------|---|--------|--------------|--------|
| # subtasks per task / # mouse clicks minimally required | 4 | 49 | 6 | 85 | 7 | 59 | 5 | 23 | 7 | 29 | 8 | 40 | 4 | 54 | 41 | 339 |
| Fictitious patient case | | | | | | | | | | | | | | | | |
| Task completion | 6/7 | (86%) | 0/7 | (0%) | 6/7 | (86%) | 4/7 | (57%) | 2/7 | (29%) | 1/7 | (14%) | 1/7 | (14%) | 2.9/7 | (41%) |
| Average # subtask completion | 3.9 | (96%) | 5.0 | (83%) | 6.9 | (98%) | 4.4 | (89%) | 5.1 | (73%) | 5.6 | (70%) | 3 | (75%) | 33.8 | (82%) |
| Average # mouse clicks needed for completed subtasks | 70 | (151%) | 72 | (114%) | 63 | (109%) | 32 | (164%) | 34 | (240%) | 37 | (131%) | 72 | (182%) | 381 | (156%) |
| Real patient case | | | | | | | | | | | | | | | | |
| Task completion | 4/7 | (57%) | 0/7 | (0%) | 2/7 | (29%) | 2/7 | (29%) | 0/7 | (0%) | 0/7 | (0%) | 1/7 | (14%) | 1.3 | (18%) |
| Average # subtask completion | 3.0 | (75%) | 3.7 | (62%) | 4.3 | (61%) | 3.6 | (71%) | 3.6 | (51%) | 4.7 | (59%) | 3 | (75%) | 25.9 | (63%) |
| Average # mouse clicks needed for completed subtasks | 59 | (145%) | 61 | (108%) | 30 | (139%) | 22 | (131%) | 17 | (194%) | 25 | (117%) | 46 | (115%) | 261 | (136%) |
| TOTAL | | | | | | | | | | | | | | | | |
| Task completion | 10/14 | (71%) | 0/14 | (0%) | 8/14 | (57%) | 6/14 | (43%) | 2/14 | (14%) | 1/14 | (7%) | 2/14 | (14%) | 4.2/14 | (30%) |
| Average # subtask completion | 6.9 | (86%) | 8.7 | (73%) | 11.2 | (80%) | 8.0 | (80%) | 8.7 | (62%) | 10.2 | (64%) | 6 | (75%) | 59.7 | (73%) |
| Average # mouse clicks needed for completed subtasks | 134 | (148%) | 133 | (111%) | 93 | (124%) | 54 | (147%) | 51 | (217%) | 62 | (124%) | 118 | (148%) | 642 | (146%) |

predefined data entry order' was found useful to analyse the fit between the system's model of predefined data entry order and the mental model of users. Our study has several limitations. The Dutch clinical algorithm for the CR NAP was developed in 2010 and is not yet completely implemented in most CR clinics. The usability issues revealed with the MediScore CARDSS 2.0 design may likewise be caused by users' unfamiliarity with the content of the algorithm. Also we were not able to test end-users' interaction with the system in the presence of real patients in daily practice. Evaluation of MediScore CARDSS 2.0 after its implementation may therefore reveal additional usability issues.

The results of our study have been handed over to the developers to adapt the final MediScore CARDSS 2.0 system to the natural task behaviour of its users. To attain this goal, the navigational structure will be organized in a more flexible and transparent way. During the think aloud users frequently mentioned that they preferred a grouped data entry of all static, standard questionnaires before entering the dynamic data patient data (i.e., the flow through data items depends on previously entered data). After attuning the data entry order during the NAP to this preference, complete data collection will be maintained by showing users which data entry steps are finished and which steps still need additional data entry. We advise to apply cognitive methods to analyse end users' mental model (task behaviour and cognitive information processing) in the requirements analysis phase of system design. The conceptual model of the system may then be made consistent with the task behaviour of users [9;10].

Competing interests – None. **Acknowledgement** – The authors would like to thank Ité Medical B.V. for their collaboration in setting up this study. **Funding** – The study was funded by ZonMw, the Netherlands organisation for health research and development.

References

- [1] Piepoli MF, Corra U, et al: 'Secondary prevention through cardiac rehabilitation: from knowledge to implementation. A position paper from the Cardiac Rehabilitation Section of the EACPR', *Eur J Cardiovasc Prev Rehabil*, 17(1):1-17 (2010).
- [2] Rehabilitation Committee (NHS-NVVC): 'Guidelines for Cardiac Rehabilitation', The Hague, Netherlands Heart Foundation. (Available at <http://webshop.hartstichting.nl/Producten/Producten.aspx?CatID=71&pID=2665> (accesses June. 30, 2011), (2004).
- [3] Netherlands Society for Cardiology (NVVC): 'Dutch Clinical algorithm for assessment of patient needs in cardiac rehabilitation and secondary prevention', Utrecht, NVVC, English version available at <http://kik.amc.uva.nl/KIK/reports/TR2011-03.pdf> (accesses June. 30, 2011), (2010).
- [4] Shortliffe EH, Buchanan BG, Feigenbaum EA: 'Knowledge engineering for medical decision making: A review of computer-based clinical decision aids', *Proc IEEE*, 67(9):1207-24, (1979).
- [5] Wagner EH, Austin BT, Von KM. Organizing care for patients with chronic illness. *Milbank Q* 1996;74(4):511-44.
- [6] Goud R, de Keizer NF, et al: 'Effect of guideline based computerised decision support on decision making of multidisciplinary teams: cluster randomised trial in CR', *BMJ*, 338:b1440 (2009).
- [7] Boren MT, Ramey J: 'Thinking aloud: reconciling theory and practice', *IEEE Transactions on Professional Communication*, 43(3):261-78, (2000).
- [8] Shiffman RN, Liaw Y, et al: 'Computer-based guideline implementation systems: a systematic review of functionality and effectiveness', *J Am Med Inform Assoc*, 6(2):104-14 (1999).
- [9] Johnson J, Henderson A: 'Conceptual models: begin by designing what to design', *Interactions*, 9(1):25-32 (2002).
- [10] Jaspers MW, Steen T, van den Bos C, Geenen M: 'The think aloud method: a guide to user interface design', *Int J Med Inform* 73(11-12):781-95, (2004).