A Rapid Usability Assessment Methodology to Support the Choice of Clinical Information Systems : A Case Study

Beuscart-Zéphir M.C.[1], Watbled L.[1], Carpentier A.M.[2], Degroisse M.[2], Alao O.[1]
[1] EVALAB, CERIM – Faculté de Médecine, 59045 Lille Cedex, France
[2]Centre Hospitalier de Denain, France

We present here an adapted methodology integrating usability engineering and early evaluation procedures to support the choice of a Clinical Information System in the context of a standard Call for Tender. We illustrate the application of this methodology with a case study. We integrated a standard 'contextual task and activity analysis' into the choice process and then drew up usability recommendations for the choice of an application. We organized a one-week on-site exhibition and test for each candidate company. During the test sessions, we performed a rapid usability assessment. The final choice of the application is strongly and positively influenced by the results of the usability assessment.

Key-words: Technology assessment; usability engineering; clinical information systems (CIS); CIS acquiring project lifecycle.

INTRODUCTION

Nowadays, a great number of hospitals have to choose and buy their Clinical Information System (CIS) or parts of this system from the industrial and commercial market. To do so, they usually follow a standard procedure: relying on Users' Requirements Analysis, they elaborate a list of desired functionalities and set a Call for Tender. Each company answering the Call provides the hospital with a detailed description of its application. Usually, the analysis and comparison of the written answers from different companies are not sufficient as a basis for an efficient choice. Hospital managers tend to compensate for this lack of knowledge about the actual usability and performances of the applications by asking the companies for demonstration sessions, and they try to visit reference sites where the applications are actually running. In spite of these efforts, the choice of a Clinical Information System or of a Hospital Electronic Patient Record (EPR) remains hap-hazardous. There is always a discrepancy between the written description of the applications and their range of possibilities and the actual dynamic activity of the target users. In a standard commercial process, it is impossible to get an early evaluation of the applications within the context of the organization. Therefore hospital managers require a helpful and efficient methodology (i) to allow a valid and realistic choice; (ii) to support the workload of the configuration of the application; (iii) to identify and support the potential necessity for partial re-engineering of the Human Computer Interface (HCI).

On the other hand, these usability and evaluation problems have been widely addressed in the past ten years [1,2]. A set of standard methods devoted to the assessment of usability features of new software applications are available [3]. We also know that early analysis and modeling of the mental processes involved in users' activity helps prevent failures and ensures better qualitative evaluation methods [4,5]. Furthermore, the integration of evaluation methodology into the Systems Design and Development Lifecycle (SDL) leads to better and dynamic assessment methodologies [6]. The integration of usability engineering methods into the SDL also allows for early and iterative usability assessment based on mock-ups and prototypes, which helps to get more usable and acceptable applications [7,8]. Therefore, a set of solutions already exists. But the problem remains that these solutions and new models have been elaborated for the **design** and **development** of new software applications: they do not apply to standard commercial procedures. However, if we look closely at the problems identified in a standard procedure of acquiring a CIS, we see that they are similar to the problems addressed in the context of Systems Design and Development Lifecycle. They concern the usability and acceptance of a medical software application in a new environment and the need to integrate evaluation methods early in the project lifecycle.

METHODS

The first two authors of this paper are usability engineers trained in cognitive psychology and ergonomics. We belong to a research lab in Medical Informatics (CERIM), and we run a Usability Lab (EVALAB) specialized in Healthcare. We were asked to address the above usability and evaluation problem. We thus adapted the existing models [7,9] to this specific situation, by integrating usability engineering methods, a phase of task and activity analysis, and early and rapid evaluation procedures into the Call for Tender process. The model of such an adaptation could be represented as follows:

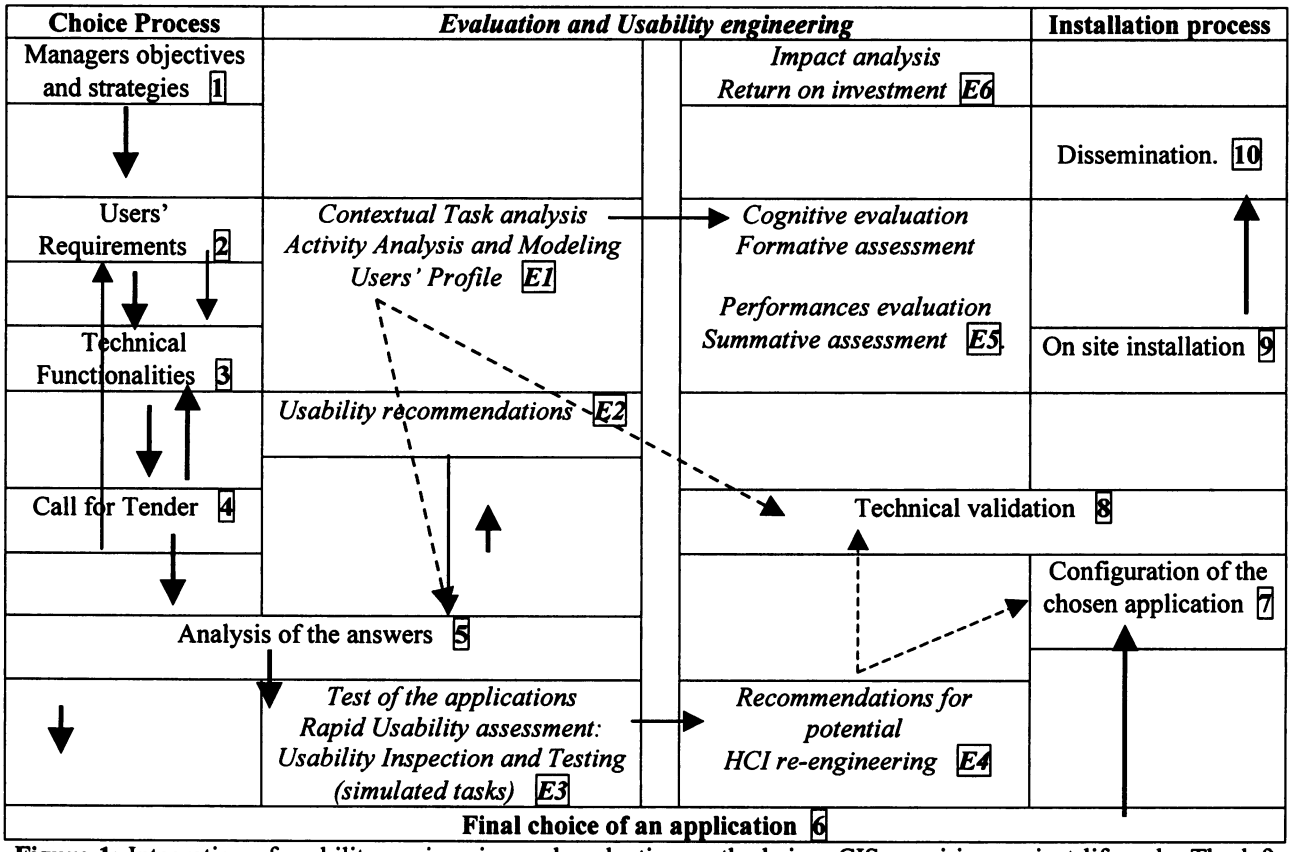

Figure 1: Integration of usability engineering and evaluation methods in a CIS acquiring project lifecycle. The left column lists the consecutive steps of the choice process (1-6). In parallel with steps 2 and 3, standard usability engineering methods take place (E1), which result in usability recommendations (E2) for the project. This in turn allows a particular analysis of the answers, from a usability point of view. A special phase of rapid usability assessment of the application is then integrated, which bears on the final choice of one application (6). If severe usability problems are identified, this phase (E3) can result in a list of recommendations for potential HCI re-engineering (E4). The right column describes the consecutive steps of the installation process (7-10). In parallel with the progressive installation, a standard assessment process is implemented (E5-E6). The main difference with standard usability engineering lifecycle in design contexts relies in the weak possibilities for re-design and iterative evaluations of the HCI.

CASE STUDY.

Based on this adapted model, we present a case study illustrating the application of this methodology to a project of acquiring a CIS in a medium sized hospital. We focus here on the choice process (steps 1-6 and *E1-E4*); the installation process is currently running.

Context of the Project.

The Denain public hospital is located in the North of France. It is a 413 bed hospital (Medicine: 80 beds; Surgery: 50 beds; Emergency: 7 beds; Maternity: 33 beds; Psychiatry: 60 beds; Convalescence and long stay: 183 beds). Around 100 physicians work in the hospital, along with 200 nurses, 160 assistant-nurses and 50 secretaries.

The project consists in the acquiring of a CIS integrating a standard medical EPR, a nursing EPR, a connected prescription system, and specialized EPR for Psychiatry, Emergency and Maternity. The users concerned by the project are all the physicians (including the pharmacists), nurses and assistant-nurses, and secretaries.

Integration of Usability Engineering Methods.

From the users' requirements document and from the list of required functionalities, we identified the domain of work concerned by the project. In agreement with the stakeholders and the users'

representatives, we selected four departments for the Activity Analysis Phase (Surgery, Emergency, one department of Medicine, and Convalescents). All the departments use the same common patient paper record, but they are under different time pressure and have different habits of work.

Contextual task analysis, activity analysis.
We used standard methods from cognitive psychology and ergonomics: (i) natural (ethnographic) observation; (ii) audio and video recorded observations with "thinking aloud" protocols; (iii) analysis of tracks of the activity and auto-facing interviews; (iv) user interviews and questionnaires. All participants in the project were given a full report of the task and activity analysis' results, which cannot be described in the limited context of this paper.

Usability recommendations.
The key points in users' activity are related to the process of medical prescription and to the management of the patient's agenda. In these domains, it is important that the application properly supports the communication between physicians, nurses and secretaries. For these main functionalities, the application must be **easy to learn** (physicians have low typing skills and are not used to computers), very **rapid**, especially where physicians and nurses are under time pressure (it takes less than two minutes to record a complex prescription in surgery), and **easy to use** by the nurses (who are used to computers and short-cuts) in order to support advanced features or tasks giving the users a sense of control.

Analysis of the answers and organization of the test period.

At the time of the writing of the Call for Tender, most of the participants in the project had made up their mind about the application they wanted for their hospital. One company (referred to as C below) had demonstrated its application several times to the hospital managers and the users' representatives, who were fairly convinced that this application was quite "ideal" for their needs.
Three companies (A,B,C) answered the Call for Tender. Two answers were better documented (A and C), but all three proposed the required functionalities. From their analysis of the answers, the usability engineers insisted that each application should be tested, and in the end, each company was asked to run a one-week on-site exhibition and test. Each company was assigned randomly to a test-week, resulting in the following order: A,B,C. The three test weeks were organized according to the same agenda.

<table>
<tr><th>Schedule / Days</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th></tr>
<tr><td>8H

12H</td><td>Installation:
1 server
3 workstations
Technical evaluation:
Professionals in charge:
Stakeholders
Technicians
Dept. of Medical Information Management
...</td><td>Discount usability inspection:
heuristic evaluation (ergonomic criteria from Bastien and Scapin [10])

3 evaluators (usability engineers)</td><td colspan="3">Demonstration and free trials. Companies demonstrators
For each user attending the demonstration and trial session, a direct usability judgment was required on three dimensions assessed with Likert scales: rapidity, ease of use (for data reading and data entering), and exhaustiveness.</td></tr>
<tr><td>12H

18h</td><td></td><td></td><td colspan="3">Discount usability testing. Usability engineers.
Simulated tasks for physicians, nurses and assistant nurses (one hour test per subject).
Observation record filled in by a trained evaluator.
For each group of functionalities tested (ex: prescriptions), a direct usability judgment on Likert scales was required at the end of the test.</td></tr>
</table>

Table 1: agenda of the test weeks; the agenda was the same for the three companies.

Subjects.

	Trials sessions			Usability testing		
	A	**B**	**C**	**A**	**B**	**C**
Physicians	9	10	16	17	11	8
Nurses	34	17	29	7	7	14
Ass-nurses	12	4	9	2	4	3
Others	6	11	15			

Table 2: number of participants in the test weeks.

A great number of users participated in the free trial and the usability testing sessions. For the nurses and assistant nurses, the users attending the usability testing sessions were always different persons. For the physicians, 3 users (out of 30) performed the test for the three applications. All the hospital departments were represented in the test.

Results.

Grading of the competing applications.
Following the three test-weeks, the participants in the project were required to draw up a report resulting in the grading of the three applications. Five different evaluation reports were thus elaborated by the following participants: (i) the hospital manager and the stakeholders (HM SH); (ii) the users' representatives (Us Rep); (iii) the usability engineers' (Us Eng); (iv) the department of Medical Information management (DMI); (v) the computer engineers of the hospital (Comp En). The results demonstrated a significant change in the attitude of the participants towards the three applications (see Table 3).

	Before the test	After the test
HM SH	C >> A #B	C ?>? A >> B
Us Rep	C >> A # B	C ?>? A >> B
Us Eng	?	A > C >> B
DMI	C > A # B	A > C >> B
Comp En	?	A > C >> B

Table 3: evolution of the grading of the three applications before and after the test weeks. Legend: >> means "very superior to", > means "superior to", and # means "not very different from".

The test sessions proved to be informative in several ways. The application B could be easily discarded. But the results of the test challenged the previous established preference towards application C. A qualitative analysis of usability characteristics was then necessary to help decide between A and C.

Results of the usability assessment.

Heuristic evaluation.
Heuristic evaluation aims at finding usability problems. The evaluators examined the applications' Graphic Users' Interfaces (GUI) according to a set of ergonomic criteria [10] and drew up a list of detailed usability problems. All those problems were rated for their severity when considering the characteristics of the target users' activity. For each application, we listed the problems recorded by all the 3 evaluators along each ergonomic criteria, and their average severity rating. We then rated each GUI using the ergonomic criteria on a five point scale, in order to allow a rapid comparison of the three applications along their respective weaknesses and good marks.
Table 4 summarizes the results of the heuristic evaluation for the three applications. Legend:
1 = severe usability problem, to be re-engineered
2 = usability problem, needs to be addressed
3 = fair, moderately satisfactory
4 = acceptable 5 = very good

Ergonomic criteria / Applications	**A**	**B**	**C**
Guidance	4	*2*	*2*
Prompting	4	3	*1*
Grouping of items	3	*2*	3
Immediate feedback	4	3	3
Legibility	3	*2*	3
Users workload	*2*	*2*	3
Users explicit control	5	4	*2*
Adaptability	*2*	*2*	4
Flexibility	*2*	*2*	4
Error management	4	*2*	3
Error protection	4	*2*	3
Quality of error messages	4	*2*	4
Error correction	3	*2*	*2*
Consistency	4	*1*	4
Significance of codes	4	3	3
Compatibility	4	*2*	*2*

Table 4: Results of the heuristic evaluation.

The heuristic evaluation shows that application C suffers from some severe usability problems especially in terms of guidance, users' control, and compatibility criteria. This means that the users could find it difficult to navigate through the numerous functionalities, and to anticipate the HCI answers to their actions. Moreover, the structure of the HCI showed low compatibility with the characteristics of the activity of the users under time pressure.
On the other hand, no tool could be considered perfect. Application A had also some serious usability flaws in terms of workload and adaptability criteria, especially for the medical prescription functionalities; which are unfortunately a key-point in users' activity.

Usability testing.
During the usability testing sessions, the users were asked to perform simple and familiar tasks with the application. The data used for the simulated tasks came from real patients' paper files and were adapted for each category of user and each department of the hospital. The results of the usability sessions largely support the heuristic evaluation findings.
For application A, physicians and nurses had great difficulties to enter any prescription properly and rapidly. Their usability rating on these functionalities were quite low, and they were reluctant to use it for their daily activity. The other simulated tasks for product A showed no major usability problems.
During the test of application C, the users kept on asking questions such as "Where am I?", and "What should I do now?". This confirms the identified guidance flaw, and leads to predict learning and training problems with the application. The nurses

found it very difficult to identify what tasks were to be done for each patient. The resulting usability rating was quite low for all those functionalities.

Final choice of an application.

The results of the test sessions affected the choice process in several ways. The hospital manager, the stakeholders and the users representatives were clearly aware of the advantages and weaknesses of each application: there was no "ideal solution". Nevertheless, they decided it would be a "GO" decision. Along with the usability engineers, they reckoned that the usability problems affecting the Application A would be easier to overcome than those of application C. Therefore, application A has been selected. The detailed usability reports were transmitted to the company A to support the demand for partial re-engineering of the prescription module.

CONCLUSION.

The adapted evaluation and usability methodology illustrated in this case-study proved to be both **efficient** and **acceptable** for all the partners.

Acceptability. The additional test phase integrated just before the choice of the product took only five weeks and led to a rapid and consensual decision. A great number of end users volunteered to participate, making the test sessions a success. None of the three candidate companies objected to the test phase, and two of them (A and C) even invested a lot of energy in this process. The cost of the test phase was less than 50000 €, including the companies' fee and the usability studies, while the overall cost of the project will probably be over 1 M€. (1€ ≅ 1$)

Efficiency. In the restricted context of the test-weeks, we had to use discount (quick and dirty) usability assessment methods. Nevertheless, the results proved to be very informative and efficiently supported the decision making process. This procedure led to a realistic choice and allowed the hospital manager to anticipate the difficulties better. Each company was sent a usability assessment report of their own test and trial sessions. From these results, two companies out of the three (A and B) decided to invest in further usability assessment and in progressive HCI re-engineering.

Side-effects. The massive participation of the users in the test and trial sessions greatly improved their knowledge of the project. At the end of the test-weeks, a physician or a nurse of the hospital would have had to be deaf and blind not to know about the project and its content.

This experience proves that it is not only possible but also efficient to integrate usability engineering and early evaluation procedures in a CIS acquiring project lifecycle. Although we need further observation to validate these results, the cost benefit ratio of this methodology already appears positive.

REFERENCES.

1. van Gennip EMSJ. & Talmon JL.
Assessment and Evaluation of Information Technologies in Medicine, 1995, vol 17, IOS press Amsterdam 1995

2. Karat J.
User-Centered Software Evaluation Methodologies
In: *Handbook of Human-Computer Interaction*, 1997, MG. Helander, TK. Landauer and PV. Prabhu eds, Elsevier, North Holland, 1997:689-703

3. Nielsen J. and Mack RL.
Usability Inspection Methods, 1994, John Wiley & Sons, New York

4. Beuscart-Zéphir MC., Anceaux F., Crinquette V., & Renard JM.
Integrating users' activity modeling in the design and assessment of hospital electronic records : the example of anesthesia
International Journal of Medical Informatics, 2001, **64** (2-3), 157-171

5. Kushniruk AW. & Patel VL.
Cognitive Evaluation of decision making processes and assessment of information technology in medicine
International journal of Medical Informatics, 1998, **51** 83-90

6. Brender, J.
Methodology for assessment of medical IT-based systems, 1997, vol.42, IOS press, Amsterdam.

7. Mayhew DJ.
The Usability Engineering Lifecycle, a practitioner's handbook for User Interface Design, 1999, Morgan Kaufmann Publishers: San Francisco.

8. Kushniruk A.
Evaluation in the design of health information systems: application of approaches emerging from usability engineering
Computers in Biology and Medicine, 2002, in press.

9. Beuscart-Zephir M.C., Sockeel P., Bossard B., Beuscart R.
Activity modeling for assessing the Usability of Telematics Applications in Healthcare
Medinfo 98, pp. 832-836 (Amsterdam IOS Press)

10. Bastien C., & Scapin DL.
Ergonomic criteria for the evaluation of Human-Computer Interfaces, 1993, Technical report 156, INRIA Rocquencourt.

Address for correspondence:
Marie-Catherine Beuscart-Zéphir.
Evalab, Cerim-Faculté de médecine, 1 Place de Verdun
F-59045 Lille Cedex, France. Tel : (+33) 3 20 62 69 70
mcbeuscart@univ-lille2.fr

Acknowledgement: this work was granted by FEDER funds from the Région Nord Pas-de-Calais