Introducing Usability Testing in the Risk Management Process in Software Development

Lindholm, Christin; Höst, Martin

Published in:
[Host publication title missing]

2013

Introducing Usability Testing in the Risk Management Process in Software Development

Christin Lindholm
Department of Computer
Lund University, Faculty of Engineering,
Lund, Sweden
christin.lindholm@cs.lth.se

Martin Höst
Department of Computer
Lund University, Faculty of Engineering,
Lund, Sweden
martin.host@cs.lth.se

***Abstract*— Human beings make errors and that is nothing that we can avoid completely. We can however lower the risk of people doing wrong in situations where, for example, medical devices are used. The overall objective of the research presented in this paper is to investigate how usability testing can contribute to software risk management process in the medical device domain. Experience has been collected from both the risk management process and usability testing in a development project of a medical device. It can be concluded that usability tests can give valuable input to the risk management process. Usability tests can indicate risks that are not identified in the risk management process and render the possibility to verify if risks with high risk value actually cause the presumed problems**

***Index Terms*—Usability, risk management, usability testing, case study, software**

I. Introduction

Medical devices and systems have an important role in today's health care and they are frequently used in different situations by different user categories. The software part in medical devices has increased over the years and plays a more and more dominant role.

A study by Walsh et al [1] shows that approximately 87% of all incidents in medical environments, where patient monitoring takes place, are due to human factors. To lower the incident rate it is thereby important to include human factors in different ways in the development process of medical devices. The purpose of this case study on a patient monitoring system is to investigate the possibilities of utilising usability testing as a contribution to the risk management process. Since risk management as well as usability are important areas in the development process of medical devices and other safety critical systems here is a need for research to investigate how these two areas can interact in a beneficial way and to implement to role of the user in different ways in the development process.

II. Background and Related Work

The user is a key player in the usability field and defined as "any human that might handle, operate and otherwise interact with a medical device though the device user interface" according to the standard IEC EN 62366, Medical Devices – Application of Usability Engineering To Medical Devices [2].

Human factors engineering (HFE) is defined [3] as the application of knowledge about human capabilities and limitations to design and development of devices, systems, tools, organisations and environments. Where as the process of human factor engineering (HFE) extends to all medical devices and has emphasis on risk management and lifecycle. There are several standards involving usability and ANSI/AAMI HE 75-2009, Human Factors Engineering – Design of Medical Devices [3] and the third edition of the medical electrical equipment standard IEC 60601-1 [4] are example of standards where usability is an integrated part of the standard.

Usability testing is regarded as a major technique for developers to use in the development process [5] in order to comply with Human Factors Engineering – Design of Medical Devices [3] and IEC 60601-1 [4]. According to Dumas and Redish [6] usability testing means focusing on the users and on how the users use the products to be productive. Usability testing is thereby a powerful method in system development based on prototyping [7]. There is a difference between the usability engineering process and the risk management process, for example, in decision making. The risk management process defines unacceptable risks, while in the usability engineering process risk are associated with usability and the design and development process for the user interface [2]. A usability engineering process focuses on all known or foreseeable hazards related to the medical user interface and not only those with unacceptable risk, like risk management process mostly do.

III. Research Method

The qualitative research in this paper is based on an empirical study in a real world setting, since process improvement activities in software engineering because of their complexity are very hard to study in isolation. The aim of qualitative research is to investigate and understand phenomena within its real life context [8], [9].

A. Objective

The overall objective of the research in this case study is to investigate how usability testing can contribute to the software risk management process in the medical device domain. More specifically the objectives are as follows:

- To investigate what type of problems and potential risks can be identified through usability testing.

- To investigate if the problems and potential risks identified through usability testing are the same problems and risks identified during a risk management process.
- To examine how the results from the usability testing can be used in the risk management process.

The objectives are investigated in a single case study at a department at a large Swedish hospital that has extensive experience in developing and maintaining medical devices, but not devices including software. The development process of a patient monitor system with an intensive care unit (ICU) as the target environment was studied during the case study. The first objective is illuminated by the results from the usability tests, the second objective by the comparison of the results from the software risk management process and the usability tests and the last objective is based on the prior findings.

B. The Case Study Context

The case study was conducted at a department at a large hospital developing and maintaining medical devices and was performed from the summer 2010 to spring 2012. The case study contains two main parts, the software risk management process and the usability testing. The focus in this paper lies on the usability testing and the conjunction between the usability testing and the software risk management process. The software risk management process is described in detail in [10].

The risk management and usability testing was carried out on a patient monitor system for monitoring a patient´s intra-cranial pressure, calculate the cerebral blood flow and present it to the medical personal on a beside monitor. The patient monitor system consists of three main parts; a) Pressure sensor placed in the patient's skull, b) Patient monitor (connected to the pressure sensor) that presents and exports blood pressure values, c) Bedside monitor, the new device which import the blood pressure values from the monitor and calculate the cerebral blood flow. The patient monitor system includes both software and hardware, although the risk management process focuses only on the software, and the usability test only on the user interface for the medical staff.

C. Case Study Process

The overall case study contains two discussion phases and three data collection phases where the first usability test is part of Phase 2 and the second usability test is part of Phase 3, see Fig. 1. The discussion phases and Phase 1 focus only on the software risk management process [10].

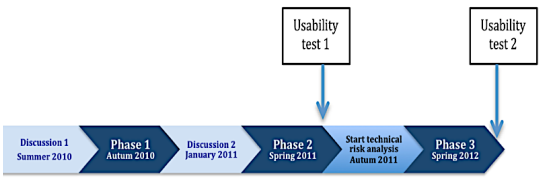

Fig. 1. Case study timeline

The detailed case study process for the usability testing part is based on the case study process described by Runeson and Höst [11]. The objectives were defined and three research questions were identified, before the preparation of the two usability tests were done and the tests were carried out with participants. The data was collected and documented during the usability tests followed by the analysis of the documentation. Results reported from the usability test were sent to the development organisation and the results from the first usability test had impact on the changes of the user interface. Feedback from the development organisation was then received.

D. The Usability Testing

According to Nielsen [12] is it enough to run a usability test with a small number of users (4 ± 1), and Virzi [13] suggests that a usability test involving 5 participants can yield 80% of the possible findings. In this case study 4 test users participated in the first usability test, and 5 test users in the second. In the second usability test there were 5 test users available at the usability test occasion so it was decided to engage all five in the test. The test users in the first usability test were 2 nurses and 2 enrolled nurses in the age of 26-33. The selection criteria were that they had not worked with the tested system before, but were experienced in using monitor equipment. Gamer et al [14] describe that novices are important test persons since they encounters most of the serious problems and also make the most errors. The test users in the second usability test were selected from the same premises and consisted of 3 nurses and 2 enrolled nurses in the age of 31-51. The test users for both test occasions were selected by the development organisation and the test facilitator and the observer prepared the test scenarios for the two usability tests. The aim of the performed usability tests was to find as many as possible of the most problematic problems. Different test scenarios were designed, for example that the test person should identify different curves on the screen, make notes and react to alarms. The usability tests were held at the intensive care unit at the hospital and preformed on the bedside monitor connected to a patient monitor. However the patient monitor was not connected to any patient. Instead the values were simulated with a simulator.

The test method that was used was "Active intervention" [6]. However, the test person was also encouraged to think out laud [12], [15] while using the system and verbalise her thoughts.

The test facilitator gave the test persons simple instructions about what to do, and encouraged them to express their thoughts. The test facilitator asked for example the test person to explain what she would do next and why. Each usability test session lasted for about 30 minutes and after each session the facilitator and observer took a few minutes to summarise and write down the things that struck them as complement to the log written during the test session. The first usability test identified 12 usability problems and the second usability test identified 16 usability problems. After each usability test the problems were presented in a test rapport supplemented with change suggestions. These reports were sent to the development organisation. The usability problems

and the change suggestions were discussed by the development organisation and resulted in a major change of the user interface after the first usability test.

E. The Software Risk Management Process

The software risk management process, applied in this case study focuses on user risks and the first three first steps of the risk management process, i.e. risk identification, risk analysis and risk planning. There were in total 15 risk meetings held from the summer 2010 until the spring 2012. Three different groups of participants were represented at the meetings; a) The intended users with special domain knowledge, e.g. physicians and nurses, b) the development organisation, e.g. medical device expert, risk analysis supervisor, and developers, and c) the researchers, e.g. process experts and technical experts from academia. At least two representatives from each group of participants were present at the meetings. For the first step, the risk identification, scenarios were chosen by the development organisation to be the main risk identification source. A scenario was defined as a chain of events, with a cause-effect relationship that describes a realistic diagnosis sequence during normal use. The risks were identified through brainstorming on each scenario and where all participants suggested possible risks connected to the specific scenario. Then each risk was assessed separately according to probability, severity and detectability. Scales predefined by the Swedish national board of health and welfare was used for probability and severity assessment and all identified risks were documented during the meetings. Both scales are four-graded (1-4). The risk value, R, was calculated for each risk by multiplying the probability, P, by the given figure for severity, S, i.e., $R = P \times S$. The highest risk value a risk in this study can have is $R = 4 \times 4 = 16$. Detectability was assessed according to the three following statements "if the fault (hazard) **always** could be detected before a severe situation occurred", "if the fault (hazard) **sometimes** could be detected" or "if the fault (hazard) **never** could be detected".

F. Data Collection and Analysis

The data collection from the software risk management process was carried out through active observations by the researchers at fifteen risk meetings. All the risks were documented during the meeting in Excel by the development organisation. In total 225 risks were identified out of which 25 risks were removed since they were not regarded as actual risks after more careful consideration.

The data collection from the usability tests was made at the usability test sessions at the intensive care unit, the first test in May 2011 and the second in May 2012. Each usability test took approximately 30 minutes and the observer logged all the actions. All observations were written down during the sessions and then transcribed on computer, resulting in reports on test results. The transcribed results were used by the facilitator and observer to identify the usability problems. First the facilitator and the observer identified the problems separately, then they compared the results, discussed the identified problems and then discussions resulted in one list of usability problems for each test. The lists were complemented with change suggestions and resulted in written test reports that were sent to the development organisation.

The data in this study have been collected from the risk documentation from the risk meetings and the documented test results from the usability test sessions.

Each of the 26 identified usability problems were sorted into three different categories based on what functionality or feature each user problem was connected to. The three categories are:

A: Alarm, problems connected to the alarm function

C: Comments, problems connected to the commenting function

D: Different usability problems, problems connected to different functions.

The usability problems in each category was given a unique identifier, for example "A2-1,2", where A stands for the category A, 2 is a serial number and 1 means registered in usability test 1, and 2 registered in usability test 2 . After that, the usability problems were classified by using the failure qualifiers defined in the classification of usability problems (CUP) scheme by Vilbergsdottir et al [16] shown in Table I. The usability problems can be classified differently, another way would for example be as described by Keenan et al [17] with primary categories and subcategories. It was decided to use the failure qualifiers [16] since they are straightforward, easy to understand, easy to categories after the usability test, and suitable for the user problems identified during the usability test. Each usability problem was documented with its unique identifier, a description of the usability problem, the failure qualifier, and the number of test persons that had that particular usability problem during the usability test. Each documented usability problem was then compared to each documented risk from the risk management process. For the usability problems where a corresponding problem was covered by a risk in the risk documentation, the usability problem was compiled together with the risk. To the documentation of the usability problems, the risk's unique identifier, the risk description, and the initial risk values was added. The usability problems were then sorted in two categories, those connected to an identified risk and those that were not connected to an identified risk. This procedure was repeated again after usability test 2. Recurring usability problems were especially marked and sorted to a special category. Some risks were reassessed due to actions taken to lower the risk, the new risk values was also added to the corresponding usability problem.

Observer triangulation [8] was implemented by having three researchers in the risk management part and two researchers during the usability test part of the case study. All collected data was treated confidential in order to protect the participants of the study and to ensure freedom during data collection. The participants have been very cooperative and were also given the right to review the findings and give feedback.

TABLE I. FAILURE QUALIFIER BASED ON [16]

Abbre-viation	Explanation
M	**M**issing, when the test participant fails to find something in the user interface that she expected to be present.
IMM	**I**ncongruent **M**ental **M**odel, when the user interface is unclear, because it does not match the test participant´s mental model or her previous experience.
I	**I**rrelevant, when the user interface contains information/object that, while perhaps true, does not contribute to system services and is not needed
W	**W**rong, when the test participant can notice that something has gone wrong e.g. apparent programming bug.
B	**B**etter way, when the test participant suggests that something in the user interface could have been done differently.
O	**O**verlook. Sometimes the test participant is given a task but she overlooks an entity in the user interface i.e. the user does not see the existing entity or fails to realize that she is supposed to interact with it.

G. Validity

The construct validity concerns to what extent all people involved understand and use terms correctly in a consistent way. There is of course a risk that participants in the risk management or the usability study misunderstand each other and that the researchers misinterpret people in the study. We have been aware of this risk and tried to make sure that we understand the participants. The internal validity concerns to what extent causal relationships are misinterpreted or based on unknown factors. Since this type of relationship is not the focus of the study, this is not seen as a problem in the study. Concerning the reliability, the analysis is carried out by comparing identified risks and problems seen in the usability analysis. No major problems are seen with respect to this. The external validity is harder to judge since this is the first study conducted in one case setting. The results can probably be of interest for other projects where risk management is carried out for a medium sized software system. Especially, the results can be valid if the organisation is new to software development.

IV. RESULTS

A. Usability Problems

During the two usability tests, 26 different usability problems were identified in total. Two of the usability problems were the same problems identified in both tests (i.e. A3-1,2 and C1-1,2 in Table II).

The majority of the usability problems concern the commenting functionality. The user interface for the commenting functionality was changed between the two usability test, although there were still usability problems connected to the commenting functionality registered after the second usability test. Finding the function in the user interface, how to add a comment, the use of medical staff identification, and how to save a comment are some examples of usability problems registered with respect to the commenting functionality. Two of the users actually pressed the wrong button when trying to save a comment and then believed that they had saved it. Here it can be noticed that the physicians at the intensive care unit find it highly desirable that the all categories of medical staff adds comments in the system.

Usability problems were also found for the alarm functionality, such as how to interpret the alarm and how to reset the alarm when it started to signal. The alarm function is vital and since two of the users, one in each test round did not notice the alarm at all, the functionality was highlighted in the development organisation. There was an in-depth discussion about adding acoustic alarm as a complement to the visual alarm, but the final decision was to avoid acoustic alarm due to the risk of alarm fatigue. Compared to an ECG-machine for heart surveillance, an alarm on the bedside monitor is not equally unguent to attend to, which also favoured having only a visual alarm. The visual alarm functionality was however redesigned after the first usability test.

The usability problems are classified according to what type of problem as presented in Table I. The classification shows that IMM - **I**ncongruent **M**ental **M**odel and O - **O**verlook are the dominant types of usability problems in this case, see Fig. 2. An IMM problem is when the user interface is unclear, because it does not match the test participant´s mental model or her previous experience, and an O problem is when the user does not see the existing entity or fails to realize that she is supposed to interact with it.

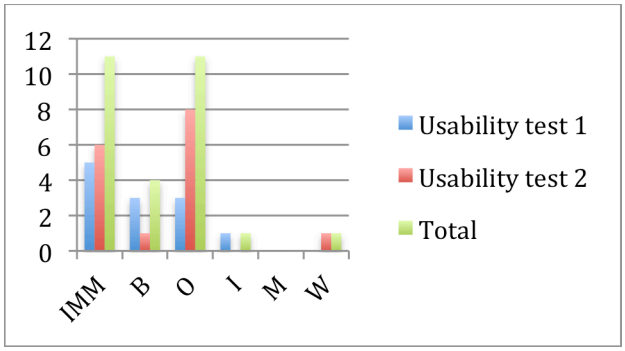

Fig. 2. Number of usability problems per failure qualifier

It was quite obvious that the design of the alarm function and the commenting function was not in compliance with the users' previous experiences of these types of functions. Concerning the commenting function, the users had trouble seeing the existing entity (failure qualifier O), more precisely where and how to add a comment in the system and also the use of standard comments. The users came with some suggestions about how things could be done better (failure qualifier B), and when it concerned the user manual they were unanimous regarding append more images in the manual.

B. Usability Problems versus Risks

If we compare the usability problems identified through the usability testing and the risks identified during the risk management process it is found that 11 out of the 26 usability problems had been identified as risks during the risk meetings. Out of these 11 usability problems, 3 usability problems were uniquely identified during the first usability test, 6 were uniquely identified during the second usability test, and two usability problems were identified during both usability tests, see Table II. For the risks connected to the usability problems, severity, *S*, and probability, *P*, were estimated at the risk meetings, and the risk values, *R*, were calculated for each risk, $R = P \times S$. The highest risk value a risk in this study can have is $R = 4 \times 4 = 16$. Initially 2 of these risks had low risk values (R = 2, R = 4) and 4 had high risk values (R = 8, R = 9, R = 12, R =16) and 4 risks was given risk value zero, since the severity and probability for these risks was regarded very low. The second last usability problem in Table II, i.e. D2 was identified as a risk but was regarded as a strict technical risk, so the estimation and handling of this risk was postponed to a later technical risk meetings. Risk value 8 was set by the development organisation as the limit for high risk-value All identified risks with risk value 8 or above were handled and dealt with. The two usability problems, A2-1 and A3-1,2 scored high risk values concerning the users' perception of the alarm function. The alarm functionality and its related risks rendered most discussions during risk meetings. There were different opinions among the participants, but the discussions resulted in a major redesign of the alarm functionality and the development organisation together with the users finally decided not to implement an acoustic alarm. The risk with only using visual alarm was put as residual risk. However the usability problem, C1-1,2, adding a comment in the system, was a problematic function for all four users in the first usability test and all five users in the second usability test. It was given a relativity low risk value (R = 4) during the first part of the risk management process with no redesign as consequence. The probability was set to 4, which corresponds well with the result of the usability test and the severity was set to 1, "discomfort or minor injury to the patient". The physicians' great desire that the commenting function should be widely used by all the medical staff is not caught in the risk management process in the beginning. As a result of the first usability test, that pinpointed the problems with this functionality, a redesign was decided. The risk was reassessed and the probability value was lowered to 2, which resulted in a new lower risk value, R = 2. This was too optimistic since the second usability test performed on the new design and after the reassessment, showed that it still was a problematic function for all the test users. There were two more risks connected to usability problems that were reassessed (i.e. C9-2 and D6-2) due to actions taken to lower these risks. The reassessment resulted in low risk values (R=3, R=4) presented in Table II with italic, underlined figures. Interesting to notice here, is that the usability problems C9-2 and D6-2 were found by all of the participants during the second usability test. The action taken had not lowered the risks to extent as expected by the risk management group.

There were 15 usability problems that were **not** caught in the risk management process, 6 of them were identified in the first usability test, and 9 in the second usability test. These problems are presented in Table III. All of these fifteen usability problems, found in the second usability test were all new problems, not found in the first usability test. It was mainly usability problems concerning the commenting function that was not documented as risks during the risk management process. There were also several problems in the D category, with problems for example regarding the touch screen, the user interface, and user manual, that was caught in the usability tests and that were not documented as risks during the risk management process. However, the problem that the users have with finding the commenting function was identified as a risk but with low risk value. On the other hand, the risk that the users would not find their way through the commenting function when for example entering text and saving the added comment was not identified as a risk.

TABLE II. USABILITY PROBLEMS CONNECTED TO IDENTIFIED RISKS

Id. usability problem	Description usability problem	Failure qualifier	Risk		
			S	P	R
A2-1	The user does not know the cause of the alarm, does not know how to interpret the alarm.	IMM	4 *3*	2 *2*	8 *6*
A3-1,2	A visual alarm was simulated; the user did not notice the alarm. The user does not see the entity.	O	4 -	3 -	12 *0*
C1-1,2	The user is given the task to add a comment in the system, but the user have trouble to find the way to do it. The user failed to find the way even if the entity existed.	O	1 *1*	4 *2*	4 *2*
C9-2	The users did not perceive the button to press for changing time for the comment. The user does not see the entity.	O	3 *3*	3 *1*	9 *3*
D2-1	The change of the graphs due to user action is unclear to the user.	IMM	-	-	-
D5-2	The users did not perceive that the graphical scales where changed. The user does not see the entity.	O	4 *4*	4 *1*	16 4

For example two users thought that they had saved the comment they had entered in the system but they had not, since they pressed the wrong button for saving the comment. For four of the usability problems found in the second usability test, all five test users noticed them. Two of these faults concern the commenting function, i.e. C10-2 and C12-2, and two concern different interface functions, i.e. D7-2 and D9-2. After the first usability test the development organisation took all the found problems under consideration and the discussions led to actions regarding all the problems except replacing the touch-screen to a more sensitive one. The system was updated and a new version was released before the second usability test.

If we look at the different types of usability problem versus identified risks, we find that the dominant class is IMM-Incongruent Mental Model, when the user interface is unclear to the user, see Fig. 3. There is a slight dominance of problems that were not at all highlighted in the risk management process and those that was highlighted in the risk management process but had got a low risk value, which meant that no action was taken according to them, although they proved to be a problem for the users. For the category O – Overlook, there is a slight dominance of problems highlighted in the risk management process which indicates that it is easier to identify items the users may overlook For the usability problems classified as B – **B**etter way, I – **I**rrelevant, and W – **W**rong,, there were more usability problems identified as risks than not identified as risks. The users did not find anything missing that they had expected to be there (i.e. M – Missing) when they took part in the usability test.

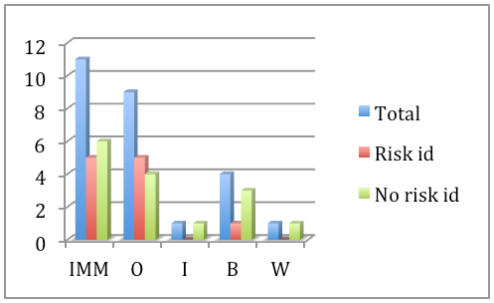

Fig. 3. Usability problems per failure versus identified as risks or not

V. Discussion and Conclusion

In the usability test there were two functionalities that generated most of the usability problems, it was the commenting function and the alarm function. The medical device used before (i.e. a sensor connected to the patient and a printer) and was replaced by the new bedside monitor had none of these two functionalities, so they were new to the users for this kind of monitoring. The users are used to alarms, but then mostly a combination of acoustic and visual alarms, and not only visual alarms as in this case. Commenting functions are available in some medical devices, for example in continuous EEG (Electroencephalography) monitoring of patients in intensive care, but it is not a common functionality in medical devices for monitoring overall.

One of the dominating types of usability problems found during the usability test were IMM, **I**ncongruent **M**ental **M**odel, when the user interface is unclear, because it does not match the test participant's mental model or her previous experience. For some of the functionality causing IMM problems, the users lacked experience, and for some of the functionality the users' mental models were not the same as the developers'. The users expect the user interface to follow their logic and not the software's or the developers' logic, so when there is a mismatch it will show as a problem. Since active intervention was used during the test, it gave the test facilitator and observer a good understanding of the users' problems and also their mental model of the product. The other dominating type of usability problems were O, **O**verlook, the users do not see the existing entity or fails to realise that they are supposed to interact with it. The users and developers perceive things differently. Things that are obvious for the developers are not even noticed by the users, and the users see and interact with the medical device in their context and on the basis of their domain knowledge. In this case, user representatives have been part of the development process and the risk management process but there have not been representatives from the whole user spectra. Since the users are novices to the tested system and lack the experience from it, it may have affected their self-confidence and made them more critical and inclined to suggest improvements. If we look at the type of usability problems according to risk, for those problems both found in the risk analysis and during usability tests there was a dominance of the Overlook category. The participants in the risk management process identified more risks with users interacting and finding the functionality than risks concerning the users' mental model of the functionality and the workflow. The Overlook problems are probably more concrete and easier to imagine for the developers when looking at the user scenarios.

There were 15 usability problems found during the usability tests that were not identified as risks. Several of these usability problems imply risk and should be handled in the risk management process, especially the four usability problems that were found by all the users. If we then consider usability problems identified during both risk analysis and usability test there were four problems with high risk value, so there seemed to be a good match between high risk value and problems for the users. However not a total match because one of the risks (connected to usability problem A3-1, 2) had a high risk value but only one user in each usability test had that problem. The probability value for that specific risk was set quite high during the risk analysis and according the results from the usability tests it should maybe not have been give such a high value after all. There were also two identified risks with high risk values in the risk analysis but they were not identified as usability problems in the first usability test. The design was changed without regarding the usability test results and these changes generated usability problems (C9-2 and D5-2) for all the users in usability test two. This indicates the need to verify if an identified risk really is a problem to the users before any changes are made.

TABLE III. USABILITY PROBLEMS NOT CONNECTED TO IDENTIFIED RISK

Id. usability problem	Description usability problem	Failure qualifier	Users that found the problem	Comment
C2-1	A text label on a button is not understood by the user, so the user does not press the button to perform the given task.	IMM	3	The fourth marked spontaneously that the button should have a better text label
C10-2	The standard comments in the system is not noticed by the users and therefore not used. The user does not see the entity.	O	5	
D1-1	The user did not notice the text information.	I	3 (4)	1 user did see the text information after a while and 3 did not see it at al.
D7-2	Users had trouble pressing the button "Back" due to its position on the screen.	W	5	

It can be concluded that usability tests can give valuable input to the risk management process. Usability tests can indicate risks that are not identified in the risk management process and give a possibility to verify if risks with high risk value actually cause the presumed problems. It is also possible to capture "problem functionality" e.g. for functionality that is new or unknown to the user. Usability testing also catches problems that are good risk candidates, where the functionality is unclear to the users and where the developers and the users have different mental models. Timing is important when it comes to usability testing connected to the risk management process. The time must be right, so no changes are made only based on the risks, before the usability test is performed. The usability tests can for example verify that a risk with a high risk-value actually is a problem for the users before any changes are made. Risk values are assumptions so if they can be identified in additional ways before any action is taken, effort and time can be saved due to the avoidance of unnecessary changes.

REFERENCES

[1] T. Walsh, P. C. W. Beatty, "Human factors error and patient monitoring", Physiological Measurement, vol. 23, no. 3, 2002, pp. 111-132.

[2] IEC EN 62366, Medical Devices – Application of Usability Engineering To Medical Devices, http://www.iso.org/iso/catalogue_detail.htm?csnumber=38594, 12/17/12

[3] ANSI/AAMI HE 75-2009, Human Factors Engineering – Design of Medical Devices, www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/HumanFactors/UCM290561.pdf, 12/17/12

[4] IEC60601-1 http://www.iso.org/iso/catalogue_detail.htm?csnumber=45605, 12/17/12

[5] J. Daniels, S. Fels, A. Kushniruk, J. Lim, J.M. Ansermino, "A Framework For Evaluating Usability of Clinical Monitoring Technology", Journal of Clinical Monitoring and Computing, vol. 21, 2007, pp. 323-330.

[6] J. S. Dumas, J. C. Redish, A Practical Guide to Usability Testing, Revised edition, Intellect Books, Exeter, England, 1999.

[7] A. Kushniruk, "Evaluation in the design of health information systems: application of approaches emerging from usability engineering", Computers in Biology and Medicine, vol. 32, 2002, pp. 141-149.

[8] C. Robson, Real World Research, 2nd edition, Blackwell Publishers Ltd, Oxford, 2002.

[9] R. K. Yin, Case Study Research Design and Methods, 3rd edition, Sage, Thousand Oaks, Californian, 2003.

[10] C. Lindholm J. Pedersen Notander, M. Höst "A Case Study on Software Risk Analysis in Medical Device Development", In proceedings of Software Quality Days 2012, Vienna, Austria, January 2012.

[11] P. Runeson, M. Höst, "Guidelines for conducting and reporting case study research in software engineering", Empirical Software Engineering, vol. 14 Issue 2, 2008, pp. 131-164.

[12] J. Nielsen, "The Usability Engineering Life Cycle", Computer, vol. 25, Issue 3, 1992, pp. 12-22.

[13] R. A. Virzi, "Refining the test phase of usability evaluation: How many subjects is enough?" Human Factors, vol. 34, Issue 4, 1992, pp. 457-471.

[14] K. Garmer, E. Liljegren, A-L. Osvalder, S. Dahlman, "Application of usability testing to the development of medical equipment. Usability testing of a frequently used infusion pump and a new user interface for an infusion pump developed with a Human Factors approach", International Journal of Industrial Ergonomics, vol. 29, 2002, pp.145-159.

[15] H. Sharp, Y. Rogers, J. Preece, Interaction Design: Beyond Human-Computer Interaction, 2nd edition, John Whiley & Sons, Ltd, West Sussex, 2007.

[16] S.G. Vilbergsdottir, E. T. Hvannberg, E. L-C. Law, "Classification of usability problems (CUP) scheme augmentation and exploitation", in Proceedings of NordiCHI 2006 (Oslo, October 2006), ACM Press, 281-290.

[17] S. L. Keenan, H. R. Hartson, D. G. Kafura, R. S. Schulman, "The usability problem taxonomy: A framework for classification and analysis", Empirical Software Engineering, vol. 4, 1999, pp. 71-104, 1999