CHALMERS

Design of a usable game based, skill training application helping the screening personnel to refine measuring of the abdominal aortic diameter in ultrasound images

Master of Science Thesis in Interaction Design

WERONIKA TANCREDI

Chalmers University of Technology
University of Gothenburg
Department of Computer Science and Engineering
Göteborg, Sweden, July 2011

Design of a usable game based, skill training application helping the screening personnel to refine measuring of the abdominal aortic diameter in ultrasound images

WERONIKA TANCREDI

Examiner: OLOF TORGERSSON

Chalmers University of Technology
University of Gothenburg
Department of Computer Science and Engineering
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Department of Computer Science and Engineering
Göteborg, Sweden July 2011

Although the screening of abdominal aortic diameter helps to identify men with abdominal aortic aneurysm and saves lives, there is need to coordinate and synchronize screening personnel's way to work. The goal of this master's thesis is to design a game based skill training application that could give the screening personnel an additional opportunity to refine measuring of abdominal aortic diameter in ultrasound images.

The design work follows the steps of the Goal Directed design process. Consequently, the design activities are divided into six phases: the Research, Modelling, Requirements Definition, Framework Definition, Refinement and Development support. The design process described in this master's thesis finishes with the usability testing of an interactive prototype. Evaluation of the design was conducted with end-users by studying their subjective ratings and performance on given tasks.

The overall results of the usability testing show that the interactive prototype of the skill training application is not yet fully usable. Consequently, further improvement of the interface design is needed.

The identified usability issues and rich qualitative and quantitative material about the interaction between test participants and the interface, collected during the usability testing, can guide the next design iteration and lead to more usable design.

Contents

1. Introduction

The first chapter of this report describes a general background of this master's thesis. The abdominal aortic aneurysm (AAA) and the screening project of Region Västra Götaland (VGR) are introduced. Further, the possibilities of skill training with games are presented. Additionally the purpose and the goal of the study are described. Eventually, the limitations of this master's thesis are specified.

1.1 Background

Abdominal aortic aneurysm (AAA) is a widening of the aorta in the abdomen area that is over 30 mm in diameter (Bergqvist, 2007). The aneurysm weakens the wall of the aorta and leads to aorta rupture and death (Bergqvist, 2007). Ruptured AAA causes death of about 1-2% of men over 60 years old (Johansson & Swedenborg, 1986; Pleumeekers et al., 1998). It is estimated that in Sweden the rupture of the aorta leads to nearly one thousand events of death each year (Wanhainen, 2007). Detecting the condition early helps to prevent rupture and death (Pleumeekers et al., 1998; Singh, 1998). If the widening is exceeding 50 mm, the AAA can be surgically repaired, if the aortic diameter is narrower than 50 mm but wider than 30 mm, the patient is controlled regularly (Bergqvist, 2007). The aortic diameter is measured by ultrasonography which is an accepted and reliable method used in screening (Lindholt et al., 1999). To detect the aorta widening early and prevent death the screening is offered to all men.

Screening is defined as “(...) the systematic application of a test or enquiry to identify individuals at sufficient risk of a specific disorder to warrant further investigation or direct preventive action, amongst persons who have not sought medical attention on account of symptoms of that disorder.” (Wald, 2001:1).

In the Region Västra Götaland (VGR) screening of the aortic abdominal diameter was introduced in 2008 and it is offered to all men, at the age of 65 (Wiel-Hagberg, 2008). This systematic, large scale measurement of aortic diameter is expected to identify men with AAA and save lives. However studies performed in Sahlgrenska University Hospital show that the measurement of aortic diameter differs between screening performers with around 5-7% (Dijnér, 2010). The same study indicates that internal training and knowing methodology contributes to more accurate measurement. Patients with identified AAA are controlled regularly and first when the aortic diameter is wider than 50 mm surgery is considered. It is therefore essential that measurement of aneurysmal aortas is as accurate as possible. At the same time it is known that measuring wide aortic diameter can be more challenging and contributes to higher variation in measurement (Dijnér, 2010). Consequently, it was decided that there is need to coordinate and synchronize the screening personnel's way to work by providing an additional opportunity to train and improve measurement of abdominal aortic diameter.

Games are proven to be an effective solution for improving work-related skills (Sitzmann, 2010). A study shows that adult trainees who received a digital game based instruction gained 14 percent higher skill-based knowledge levels than trainees instructed in a traditional way (e.g. by completing assignments) (Sitzmann, 2010).

Similarly, Totty (2005) states that adults learn better and retain more if a learning course incorporates game elements. More compelling and fun company education programs are believed to make learning more effective (Totty, 2005; Yaman, 2001). Even simple, unsophisticated games motivate players and make them compete and engage in learning (Totty, 2005, Sitzmann, 2010).

Yaman (2001) states that positive emotions, excitement and stress relief that occur while playing games maximises learning. People remember information tied to strong emotion and therefore learning through games is believed to be more successful (Yaman, 2001). Additionally a game can be played repeatedly and consequently give players possibility to master the skill that is being taught (Brothers, 2007; Sitzmann, 2010).

Furthermore, according to Whitton (2010) the strength of games as a training tool is the fact that they provide a unique possibility to make mistakes in a safe environment and learn from these errors. In addition, as the results achieved in a game are monitored by the system game-based training gives trainees and stakeholders possibility to watch and assess the skills that were taught (Bergeron, 2006).

1.2 Purpose

The purpose of this study was to explore if a game based on-line training application could be a viable and acceptable solution helping the screening personnel to train measuring abdominal aortic diameter in ultrasound images.

1.3 Goal

The goal of this master's thesis was to design a usable game based on-line skill training application helping the screening personnel to train measuring abdominal aortic diameter in ultrasound images. The Goal Directed design process was applied. Eventually, usability of a high fidelity interactive prototype was evaluated with users by studying their subjective ratings and performance on given tasks.

1.4 Limitations

This study does not examine if repetitive use of the designed game based on-line skill training application leads to more accurate measurement among the screening personnel or quicker learning of the work practice.

1.5 Structure of the report

The second chapter of this master's thesis introduces the theoretical background of the design process: user centred design and educational theories.

The third chapter presents the method and follows the steps of Goal Directed Design process. The activities of the Research, Modelling, Requirements Definition, Framework Definition, Refinement and Development support phases of the Goal Directed design process are described.

The outline of the results chapter follows the iterations of the Goal Directed process. Initially, the outcome of the Research phase is shown, followed by the results of the Modelling phase, then the findings of the Requirements Definition, Framework Definition, Refinement phases are described, finally the outcome of the Development support phase is presented to the reader. The findings of the final usability testing of the interactive prototype with end-users are summarised in the results chapter as well.

This master's thesis ends with a discussion of applied methodology. Further, it is discussed whether the results of the usability test of the interactive prototype meet previously established usability goals. The conclusion summarises the results and contribution of the entire design process.

2. Theory

2.1 User centred design

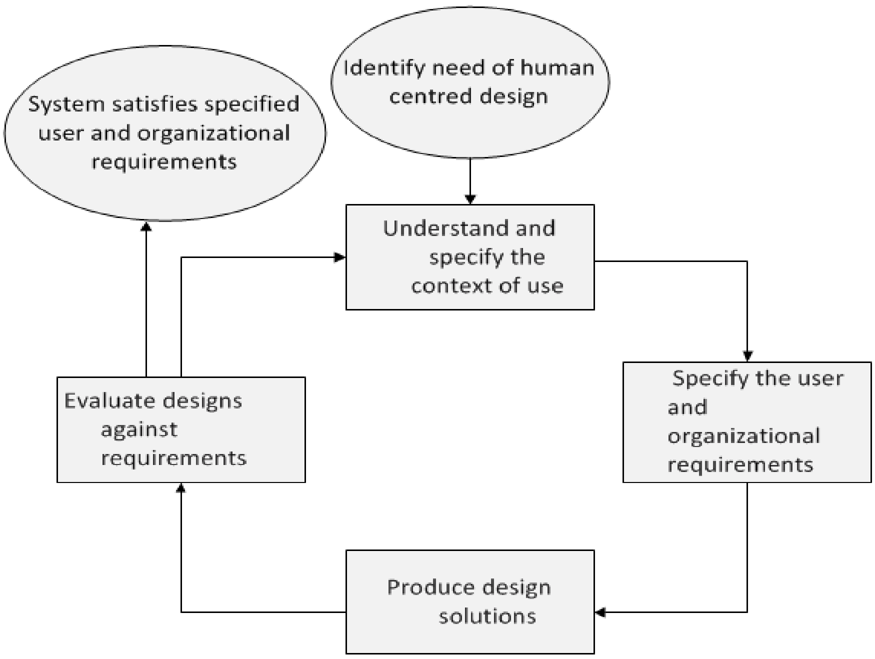
Successful products and systems must be developed with focus on usability (Beyer & Holtzblatt, 1998; Maguire, 2001; Cooper, 2007). Only then the training and support costs can be reduced and user satisfaction and productivity can be improved (Maguire, 2001; Jokela, 2008; Cooper, 2007). When user centred design is applied and when end users and their goals influence the development of software, the final product supports the users instead of constraining them (Sharp et al, 2007).
User centred design for interactive system is “an approach to interactive system development that aims to make systems usable.”(ISO 13407 (1999) cited in Jokela, 2008, p.170). To achieve usable products four user centred design principles are recommended:

- *User involvement* —the active involvement of users and a clear understanding of user and task requirements.
- *Allocation of function* —an appropriate allocation of functions between users and technology. The decision of what the system does and what actions are left to the user should be based on analysis of capabilities of both human and technology; user's well-being should be taken into consideration.
- *Iteration* —iteration of design solutions.
- *Multi-disciplinary design* —different skills are required in User centred design.

Additionally the same Standard provides four activities that are necessary for User centred design:

- Understand and specify context of use
- Specify the user and organizational requirements
- Produce design solutions
- Evaluate designs against requirements

Figure 1. User centred design according to ISO 13407 (1999)



According to ISO 13407 (1999), user centred design starts with planning of design activities. In this first phase the resources for all activities are allocated. The design process is completed when the system or product satisfies the previously specified user and organizational requirements.

The ISO 13407 standard provides an overview and a general guidance on human centred activities but it does not provide details concerning usability methods that should be used within each of four activities (Jokela, 2008). The game based training application for the screening personnel is a novel product and therefore it was decided that a well structured process with clearly characterized methods should be followed. Consequently Goal Directed Design Process (Cooper, 2007) was applied.
Goal Directed Design is a user centred methodology providing guidance throughout the product development process. Designers following Goal Directed Design establish what goals and needs users have, create a model of users and finally develop a solution satisfying users' needs. The steps taken within the Goal Directed Design process are presented in the next chapter.

2.2 Educational theories

Quintana et al. (2003) present three approaches for designing tools for learning: the behaviourist approach, the information processing approach and the social constructivist approach.
Behaviourism assumes that it is possible to shape and influence people's behaviour by rewarding desired and punishing the unwanted or incorrect behaviour (Quintana et al., 2003). Software designed according to this approach includes immediate feedback and constant measurement and monitoring of learning progress (Quintana et al., 2003). Examples of such applications, presented by Quintana et al (2003), are Reader Rabbit, Math Blasters, and Typing Tutor for teaching reading, arithmetic, and typing respectively. The users of these educational games perform activities and are either rewarded or punished (Quintana et al., 2003). Game points are given each time a task is done correctly, the incorrect activity results in ”negative action on the learner's game character” (Quintana et al., 2003:827).
The behaviourist approach is recommended for training specific skills that have well defined answers.

Information processing is the second approach to developing teaching tools. The method builds on models of cognition (Quintana et al., 2003). Information processing assumes that the knowledge necessary for performing activities can be represented and formulated (Quintana et al., 2003). Applications that follow the information processing approach often incorporate intelligent tutoring systems, ITSs (Quintana et al., 2003). ITSs build on two models: a certain task model and a learner's cognition model, the system then monitors the learner's work and provides help and assistance if necessary (Quintana et al., 2003).
An example of an intelligent tutoring system is LISPITS, a program teaching LISP programming that models the steps needed to write a LISP program. It then compares the actual steps that the student takes with this model (Urban-Lurain, 1996). LIPSITS immediately informs the learner of an error and explains how the student should continue to solve the task (Corbet and Anderson, 1988).

The third approach to educational software development presented by Quintana et al (2003) is *social constructivism*. This approach is recommended for learning wider, open-ended work practices (Quintana et al., 2003). The social constructivist approach includes both technology and human mentors or collaborators (Quintana et al., 2003). This model assumes that learning is an active process where learners need context, discourse and community of peers or experts for learning (Quintana et al., 2003). An example of constructivist learning environment is collaborative computer game (Whitton, 2010).

3. Method

This third chapter starts with an overview of the Goal Directed Design process and an introduction of evaluation techniques used throughout this design process. Further, the application of the Goal Directed Design process is described. The outline of this chapter follows the six phases of the Goal Directed design process. Design activities of each of the six phases of the process are presented.
The iterative character of the Goal Directed Design influenced the content and structure of this chapter. The design solutions were created, tested and modified just to be evaluated again, in consequence the early design solutions are shown in this method chapter. Testing and modifications of these early results are seen as a part of the method and are therefore described in the current chapter.

The Goal-Directed design process

As mentioned in the previous chapter, the Goal Directed Design process was applied in this master's thesis. According to the Goal Directed Design the final behaviour and appearance of product is based on goals of users, needs of stakeholders and limitations of technology. The Goal-Directed design process consists of six phases: Research, Modelling, Requirements, Framework Definition, Refining and Supporting.

The Research phase includes qualitative ethnographic field methods. Gathering information about users employs following qualitative research techniques: stakeholder interviews, subject matter expert interviews, user interviews, user observation, literature review and examinations of competitive interfaces. The user observation was based on a qualitative field technique called *Contextual inquiry*. The approach was introduced by Hugh Beyer and Karen Holtzblatt and selected as a research method for this design process because of its extensive and exact description of practice and principles of user observation. Contextual inquiry helps designers to gather information about a work area that need to be investigated. It is recommended that between 10 and 20 users are observed as they perform their daily tasks (Beyer and Holtzblatt, 1998). The researching designer takes a role of an apprentice and focuses on learning about the structure of work, needs and desires of the selected user, the observer therefore asks the user about the details of work activities (Beyer and Holtzblatt, 1998). The user acts like a master and talks about the work while doing it. After each observation the design team interprets collected data and structures it into work models (Beyer and Holtzblatt, 1998).

The Research phase reveals behaviour patterns that are then used for building domain and user models during the Modelling phase. Domain models include diagrams showing information flow and workflow. Users with similar behaviour, attitudes, aptitudes, goals and motivations are represented as a persona.

During the Requirements phase context scenarios are developed, these generic stories describe persona's activities and desires and show how the designed product helps the persona to achieve her goals. Context scenarios aim to describe ideal user interactions and experience.

The overall structure of the product, its behaviour, elements and visual design are defined during the Framework Definition phase. In that stage the key path scenarios, validation scenarios and sketches are created. The key path scenarios describe precisely the behaviour of the interface, they show in detail how the persona does the most common tasks. The validation scenarios, on the other hand, present the infrequently performed interactions.

In the Refinement phase, the initial rough sketches are translated into screens that show the design in greater detail. The screens in form of paper prototypes are presented to both experts and users who help to identify problems with interaction. Evaluation techniques applied in this phase are introduced to the reader in the next section.

The Development support phase is the last phase, the designer answers questions of the developers as the interface is constructed.

3.1 Evaluation techniques

In the Refinement phase of this Goal Directed design process both expert based and user based evaluation techniques were applied. These formative usability evaluation methods were an important part of each design iteration. The techniques helped the designer to identify usability problems and the outcome of each testing session was influencing the design decisions concerning elements and behaviour of the interface.
Below two groups of techniques that were used through the Refinement phase of this design process are shown. Section 3.6 shows how these methods were applied in the actual design process.

3.1.1 Expert based methods

Expert based methods involve one or several human computer interaction practitioners who systematically examine the interface. The expert based methods are performed quickly and are cheap as users are not involved.

Heuristic evaluation

Heuristic evaluation is a usability inspection method where professionals with background in human computer interaction systematically judge aspects of user interface. Evaluators independently inspect elements of the user interface by comparing it to a set of guidelines (Nielsen, 2005). For instance evaluators control if the components of design inform the user about the status of the system or if they prevent the user from making errors. The goal of heuristic evaluation is to find usability problems (Nielsen, 2005).

Cognitive walkthrough

Cognitive walkthrough is a flexible method helping experts to evaluate interactive systems at all stages of the design and development process. The evaluator analyses tasks by walking through a sequence of actions composing a single task. The examination is done by answering questions regarding effects of actions, visibility of user interface elements, recognition of meaning of actions and understanding of feedback given after each action (Dix et al., 2004).

3.1.2 User based methods

User based methods involve actual or potential users. By observing users as they perform tasks on a prototype the designer can find out whether people can use the product successfully (Maguire, 2001). Assessing the usability of a product is done by collecting subjective and objective data (Tullis, 2008). Subjective data is data supplied by the test subject, self reported, in contrast objective data is performance data collected in the usability study (Tullis, 2008).

Usability testing

Usability testing, sometimes called usability evaluation, checks whether a product is usable. Usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." (ISO9241-11, 1998 cited in Hornbaek, 2006, p.82).
To assess the product's usability, usability metrics are employed. The usability testing and metrics help the designer to answer in what extent the user using a product is effective, that is if he or she can complete a task, if the user is efficient, that is how much effort is required to complete the task, or if the user is satisfied, that is if the user was happy with his or her experience when working on the task (Tullis, 2008).

In this study three usability metrics are used to evaluate the design: performance, issue-based and self-reported metrics.
Performance metrics include task success, time on task, errors and efficiency (Tullis, 2008).
Issue-based metrics are defined by Tullis (2008) as anything preventing task completion, creating confusion or producing error. Misinterpreting the content, performing wrong actions, assuming something is correct when it's not and not understanding navigation are other examples of usability issues given by Tullis (2008).
Self-reported metrics are based on what users say about their experience when using the product, self-reported metrics can include satisfaction, expectations and ease of use (Tullis, 2008).
Usability tests can be carried out as a formative or summative evaluation (Tullis, 2008). The formative evaluation is conducted before the final design is created, the formative test identifies shortcomings and usability issues, the results of the formative evaluation are recommendations of change that help the design team to improve and refine the design.
Summative evaluation is conducted when the final design is ready, the summative test assesses how well the design solution meets its objectives, that is compares the design against the target usability goals.
Subjective data can be collected immediately after the user carried a single task (post task ratings) or after the whole test (post session ratings). Both types of ratings collect data on users' perception of the product and the interaction with it (Tullis, 2008). The System Usability Scale (SUS) is a post session rating technique that was developed by John Brook in 1986 (Brook, 1996). The survey consists of ten statements, test participants select on the five point Likert scales in what degree they approve the statements. The SUS shows the test participants' subjective assessment of general usability of a system (Brook, 1996). Half of the questions are positively worded and half negatively. The ten statements included in the SUS evaluate the need for support, training, ease of learning, ease of use and complexity (Brook, 1996). Each item score is determined according to the rules. For items that are positively worded the score is the scale position minus 1. The score of negatively worded items is always 5 minus the scale position. The overall SUS score is obtained by multiplying the sum of the item score contributions by 2.5 (Tullis, 2008). Initially the usability practitioners were convinced that the statements shouldn't be evaluated separately but the outcome of the survey should be one, general rating score, ranging from 0 to 100 (Tullis, 2008). Lewis & Sauro (2009), however, encourage decomposing the overall SUS score into two components Usability and Learnability. The summed scores for Learnability questions (question four and ten) should, then be multiplied by 12.5 (Lewis & Sauro, 2009).

Paper prototyping

Snyder (2003) defines paper prototyping as “a variation of usability testing where representative users perform realistic tasks by interacting with a paper version of the interface that is manipulated by a person 'playing computer', who doesn't explain how the interface is intended to work.“
Paper prototyping is a quick and cheap way to get user feedback and identify usability problems (Snyder, 2003). After each test the problems are listed, the paper prototype is revised and tested again (Snyder, 2003).

3.2 Research phase

This section presents the first phase of the Goal Directed Design process. It shows how the observation of and interviews with users were conducted and how the collected material was then processed. Gathering requirements of the stakeholders are covered below as well.

The research phase started with identification of people who could be affected by the new training program and who could influence the requirements. Cooper (2007) defines stakeholders as those who have authority to decide about and responsibility for the product that is designed. Stakeholders are the key members of the organization ordering the design, for instance executives, managers and representatives from the development team. Sharp et al. (2006) provides a more general and wide definition of stakeholders that includes everyone influenced by the design, for example users directly handling the product, developers, recipients of the outcome of the product and managers of all teams involved in the development work.

Subject matter experts are people with long experience and extensive knowledge within the domain of the product that is designed (Cooper, 2007).

The initial interviews with stakeholders provided information about the organizational goals, the context, intended users and functionality of the training application. Afterwards, the end users were contacted and a plan for observation of and interviews with them was developed.

3.2.1 Stakeholder interviews and subject matter expert interviews

Two senior physicians from the Sahlgrenska University Hospital and a system administrator from Bild- och Funktionsregistret were suggested as stakeholders and subject matter experts and interviewed. To ensure that everyone's perspective is understood and taken into consideration the stakeholders were interviewed individually. Each interview lasted about one hour and was conducted in the Sahlgrenska Hospital and in the Bild- och Funktionsregistret, the audio was recorded.

Developing a training application requires that the designer is familiar with the current learning style of the user group (Bergeron, 2006). Consequently, the first group of questions asked to the stakeholders was about the introduction of personnel into the aorta screening project in Sahlgrenska Hospital. The second group of questions considered preliminary vision and the use of the training application. General questions about behaviour, functionality and technical limitations of the program were asked as well. Eventually stakeholders were asked to define the user group and to fill in a questionnaire.

3.2.2 Field study

Involving end-users into the design process is an essential requirement of the user centred design process (ISO, 1999) and the Goal Directed Design process. According to Cooper (2007) it is important for the designer to capture a wide range of user behaviours and therefore it is recommended to identify observation and interview subjects with diverse backgrounds. Consequently, to obtain an accurate and valid picture of end users and their working practice the screening personnel from different professional background and from various working environments were observed. Because the screening for the aortic aneurysm is conducted by three different groups of personnel it was intended to involve representatives of all three professional categories. Consequently six medical laboratory scientists, two nurses and two physicians were contacted and asked to participate in the design process of the training application.

Screening for abdominal aortic aneurysm is currently done in six locations in Region Västra Götaland: Norra Älvsborg County Hospital in Trollhättan (NÄL), Skaraborg Hospital (Kärnsjukhuset in Skövde – KSS), Southern Älvsborg Hospital in Borås (SÄS), Kungälv Hospital, Östra Hospital in Göteborg and Mölndal Hospital in Göteborg. To ensure the variety of the observation subjects, ten professionals conducting the aorta screening were observed in five different hospitals in the Region Västra Götaland. Only the personnel from Norra Älvsborg County Hospital in Trollhättan was not involved. Each observation lasted between one to two hours, the steps taken by the personnel during the examination procedure were noted. If the patient gave his permission the screen of the ultrasound machine was filmed as the diameter measurement was performed. The collected film sequences were then analysed and used as a mediating tool during the interviews, in the second visit. Due to the respect to the patients, big number of examinations and little time allocated for each patient, no questions were asked during the observation.

The second visit in each hospital lasted about one hour and focused on details and the structure of the work. It had a form of a semi-structured interview with open-ended questions and the audio was recorded. Two persons were interviewed at the same occasion. Ultrasound images of the aorta from each filmed patient were discussed. According to the Goal Directed Design process five behavioural variables should be identified, consequently questions concerning activities, attitudes, aptitudes, motivations and skills were asked. The personnel were asked to describe what steps they took when analysing each ultrasound image, what kind of problems they encountered, what skills they thought were critical for measuring right, when they knew that they measured accurately and how they continued when problems were encountered. General questions about the experience of and attitude towards web based training were asked as well. Eventually, the interviewee were asked to fill in a questionnaire collecting data about their attitude towards the continuous training in aorta screening.

Analysis of the interview data

When analysing the interview data two methods for qualitative analysis were selected: *identifying recurring patterns or themes* and c*ategorizing data* (Sharp et al., 2006). Data analysis started with tracking of patterns in the material. Since the Goal-Directed Design process assumes identification of five types of behavioural variables (activities, attitudes, aptitudes, motivations and skills) the search for recurring patterns focused on looking for themes within these behavioural variables.

The five behavioural variables were kept in mind from the beginning of the research process. The observer was trying to understand and identify users' activities, attitudes, aptitudes, motivations and skills as she watched the users perform the screening. Further, the interview questions were prepared to fill in the gaps and to gather all the missing facts but also to confirm if the conclusions drawn during the observations were accurate.

Interviews were transcribed and read repeatedly, the most representative statements describing the five variables were identified. More exactly, statements answering the following questions were selected: what the user does, how the user thinks about measuring aortic diameter and ultrasound machine, what training and motivation the user has, and what the user perceives as difficult and helpful. The data was thus divided up into elements, i.e. statements (Sharp, et al., 2006). The selected statements were then copied from the interview transcription document and pasted into a separate document called “core points” (section 4.2). In this document all statements were grouped so that statements belonging to users were presented separately. As the interviews were always conducted with two interview subjects the answers needed to be arranged so that the statements of each user could be viewed separately. This organization of data was believed to facilitate the analysis. Further, the statements collected in the “core points” document were analysed for finer patterns. In this step ten different recurring themes within the five behavioural variables were observed (see section 4.2). Each of the ten patterns were described by one sentence and referred to by alphabet letters from A to J. Further, the statements collected in “core points” document were tagged with codes (from *A* to *J*). Consequently ten categories describing *activities*, *attitudes, aptitudes*, *motivations* and *skills* were obtained.

Below an example from the core points document is shown, statements 14 and 15 of the user four (U4) are presented. <u>H</u> , <u>A</u> and <u>J</u> in the beginning of the statements are codes classifying the statements.

> <u>H</u> U4 14 när jag känner att jag knappt ser aortan så frågar jag någon annan
> *<u>AJ</u> U4 15 När man känner sig nöjd, när man ligger på gränsen så är man noggrannare, har man 16, 17 så gör man snabbt, man kan inte sitta en kvart med det.*

The analysis of the interview data was completed when ten patterns describing behavioural variables were identified. These variables were the foundation for building persona in the next step of the Goal-Directed Design process. The process of building the persona is shown in section 3.3.1.

3.2.3 Questionnaires

To collect a quantitative data about goals and work of users and stakeholders two types of questionnaires were created.

The first questionnaire was distributed to the personnel. It gathered data concerning the training and introduction of the personnel into the examination of aorta as well as the personnel's attitude toward further education within aorta measurement. Totally ten participants representing future end users from three professional groups were asked to answer the survey: two nurses, two physicians and six medical laboratory scientists.

The second questionnaire was completed by the two senior physicians previously identified as stakeholders. The survey prepared for them included similar questions to these in the personnel's questionnaire and gathered data about educational material available for the screening personnel, training of the personnel and need for further training within the aortic diameter measuring.

3.2.4 Literature review

In parallel with the interviews and observations the literature about aortic aneurysm, ultrasonography, relevant products, different types and characteristics of educational games and usability methods were reviewed. The findings of this research were then used when planning e.g. interviews with stakeholders or during the actual design.

3.2.5 Examination of competitive interfaces

Stakeholders didn't know if there was any product with exactly the same functionality to that wished for the application being designed. Therefore products providing similar functions were examined. *Picsara* (Euromed Networks, 2010) and *Siemens Syngo dynamics* (Siemens, 2010) were identified as relevant systems providing functionality similar to that in the planned training application. During the exhibition VITALIS (discussing the development of healthcare and nursing through information technology) in Gothenburg, representatives of companies introduced the functionality of both applications. Further, the demo version of Picsara was installed and explored. Due to difficulties with accessing a working version of Siemens Syngo dynamics only Picsara was explored.

Additionally, three serious games helping the healthcare professionals to develop their skills were identified as relevant for the design of the training application. *Body Simulation (METI, 2010), Cardiac Arrest! (*Mad Scientist Software, 2010) and *eDose* (METI, 2010) were analysed and the identified characteristics of these existing interfaces were helpful, in the later stage, when the new training application was actually designed. The three skill training programs were explored through the information available on the producer's web site and not through the direct interaction.

3.3 Modelling phase

The findings of the Research phase provided material for creating the persona and the workflow model. Below the steps taken when creating the persona and the workflow model are shown to the reader.

3.3.1 Creation of the Persona

According to the Goal-Directed Design process constructing personas starts in the research phase with identifying behavioural variables (Cooper, 2007). Five types of behavioural variables were considered: *activities*, that is what the user does and how often, *attitudes* – how the user thinks about the product domain and technology, *aptitudes* – what education and training the user has, his or her capability to learn, *motivations* - why the user is engaged in the activities and *skills* that is users' capabilities related to the domain and technology (Cooper, 2007).

As it was mentioned (section 3.2.2), during the research phase ten distinctive patterns within behavioural variables were identified.

Creating persona started with mapping interview subjects to behavioural variables that were identified in the research phase. For each behavioural variable a separate behavioural axis was drawn and each of the interview subjects was placed on the axis.

If all of interviewees were located in the same area of the axis then the conclusion was drawn that the identified behaviour was common and that all interview subjects follow the same behaviour trend.
The second step of constructing the persona involved identifying significant behaviour patterns. All interview subjects with similar trends created a distinctive behaviour pattern. When it was established that only one behaviour pattern for all interview subjects could be identified, the goals of interview persons were listed. In the last step, a third-person narrative was written around all identified characteristics and goals (attachment).

Observations and interviews revealed the work practise of the screening personnel. The process of building the workflow model is introduced to the reader in the next section.

3.3.2 Creation of the workflow model

According to the information processing approach to developing of the teaching tools, modelling the user's cognitive flow is necessary to design a program which can provide help and assistance for the learner. Creating of the workflow model started with analysis of the interview material. The statements collected in a “core points” document were read and all phrases explaining the work practice were marked in a bold style. A table gathering the personnel's steps, actions, difficulties and risks was created. The table helped then to create the workflow model (see appendix).

The next section shows how it was determined what capabilities of the application are important for the persona to accomplish her goals.

3.4 Requirements Definition

The Requirements Definition phase began with defining the problem and vision statements. Further, the expectations of the persona were revised and eventually, the context scenarios were created.

3.4.1 Problem and vision statements

Both statements are created to summarize user and organisational goals and present them in a formal way (Cooper, 2007). The designer should see the statements as guidelines showing the high level objectives and evaluate design decisions against the two points. The problem and vision statements are based on user and organizational goals identified in the research phase. The problem statement should clearly define the problem that both persona and organization is experiencing (Cooper, 2007). The vision statement is a high level design purpose (Cooper, 2007).

3.4.2 Identification of the persona's expectations

Attitudes, goals and statements describing functionality the users wished were reread in the interview material and a summary of persona's expectations was written down.

3.4.3 Creation of context scenarios

Three types of scenarios are used in a Goal-Directed design process (Cooper, 2007). Every scenario created during this process was a persona-based scenario describing usage of the training application by the persona. The context scenario was developed early in the design process, in this Requirements Definition phase. It contains general description of the interaction. By contrast, the key path scenario was developed later in the process, after the Design Framework was defined, the key path scenarios are therefore much more detailed(see section 3.5.5). The validation scenarios were developed late in the process. Working on validation scenarios started when a detailed mockup was built and when common use paths were stable enough to pay attention to infrequent interactions (see 3.5.6).

The persona's goals and expectations and stakeholders' desires were translated into context scenarios. These stories should show the perfect interaction between the application and the persona (Cooper, 2007). Context scenarios should describe how the product supports persona in achieving her goals (Cooper, 2007). Therefore the starting point for creating scenarios was to view expectations and goals of users and stakeholders. The workflow model of the work practice and the observation of the context of work were another important elements contributing to the context scenario.

3.5 Framework Definition

The first step of building the Framework Definition includes determining the form and posture of the application.

3.5.1 The form and posture

The behaviour of the application should be designed with respect to the way it presents itself to the user – its posture (Cooper, 2007). The main goal of the program is to help users to train measuring aorta diameter. Completing the exercises requires from users complete attention and time. Consequently the primary posture of the program is a *sovereign posture*, that manifest itself as a full-screen application with specialized panes and many controls intended for long and concentrated use. The application, however, will be accessed rarely, only twice a year, therefore several elements of *transient posture* were included as well. Functions like commenting images, viewing guidelines and results are performed quickly in relation to training measurement therefore the simple, clear, obvious interaction was used according to principles of transient design.

3.5.2 The interaction design principles

The second step in creating the definition of Framework is to establish what design principles should be the foundation of the design. Consequently it was decided that the design should be based on several interaction design principles described by Cooper (2007).
The first set of principles applied is about adjusting the interaction according to users' experience and needs. Because the application is intended for infrequent use the goal is to optimise it for easy learning. Users will access it only twice a year and should be therefore treated as first time users and get necessary support. Infrequent users forget the organization, structure and functionality of the program therefore according to Cooper (2007) an important principle is to help these infrequent users become familiar with navigation, functionality and prevent them from making errors. A single exercise includes examination of 30 patients and is estimated to be completed within one hour. The procedure of each aorta examination is identical, therefore it is expected that the user, after several measurements, will become familiar with it and become an intermediate user. Consequently the application should satisfy well two groups of users: beginners and intermediates. The aim was therefore to design a program satisfying needs of infrequent users, having characteristics of first-time users, and intermediate users. It was strived after that these more experienced users should be considered and not disturbed by supportive functionality intended for beginners.
Another important set of design principles presented by Cooper (2007) encourages designing applications that do not disturb users' flow. According to these principles the program should avoid any behaviours that disturb the flow, instead it should support users concentration and involvement. One way to support users' involvement is to design a coherent and harmonious application that gets invisible. It was intended that the application informs quietly without demanding answers when that's not necessary.
The truly harmonious design is obtained with a modeless feedback (Cooper, 2007). The intention was therefore to display information in a clear but silent and subtle way without disturbing users'

concentration. Consequently all confirmation dialogues were consciously avoided and substituted by modeless feedback.
The design was influenced by the principle of eliminating excise. All extra steps and work users need to do to interact with the program were intentionally avoided. For example the confirmation of a choice was eliminated.
Additionally it was intended to automatically support the most probable scenarios, e.g. the user should be able to see own results without requesting it.

3.5.3 Data and functional elements

After establishing what interaction design principles should be considered during the design, the data and functional elements of the future application were listed. Three basic units were identified: video sequences, exercises and results (see chapter 4). For each data element several attributes were established. Both data elements and the attributes are based on context scenarios and the workflow model.

3.5.4 Functional groups and hierarchy

In this step the identified data and functional elements were grouped to support the persona's flow within the tasks. To ensure that the grouping of elements into screens and panes was determined by the persona's mental model of work, the workflow model and description of interaction from the context scenarios were used. The final grouping is presented in chapter 4.

3.5.5 Creation of key path scenarios

The key path scenario was built upon the results of the Framework Definition. In the framework Definition phase the data and functional elements were identified. The possible operations were then listed in chronological order, arranged into screens (see table 3) and described in the key path scenario. The end of the Framework Definition phase and creation of key path scenarios were thus done in parallel.

3.5.6 Creation of validation scenarios

Infrequent interactions and needs were identified, and then a strategy to meet these requirements was developed.

3.5.7 Visual design

Using the application for approximately one hour requires conservative colour palette and visual style (Cooper, 2007; Vora, 2009). The colour scheme was selected after the observation of the working place, green was reoccurring in the clothes of the personnel, furniture and on the walls.
The intention was to make the logo relevant for the aorta screening, memorable and recognizable. It should represent or describe the measuring the aortic diameter.

When all elements of the interface were identified and grouped the first mockups were created. The next section shows how different elements of the interface were formed and evaluated.

3.6 Refinement phase

In the Refinement phase the initial sketches were translated into screens that showed the design in greater detail. The first versions of paper prototypes were created with an online wireframing tool Balsamiq Mockups, as the design became more precise a new prototype was built with Microsoft Visio 2010. To ensure that the created design supports users, four types of usability tests were conducted. The first part of this section shows how the six elements of the interface were created and refined. The second part of this section reveals how the solutions were evaluated.

3.6.1 Creation of mockups

Displaying sequences composing an exercise

The intention was to inform the user in a silent and subtle way about the number of sequences that create an exercise, what sequence is processed at the moment and how many are left. The initial idea was to present all that information graphically.

In the beginning the video sequences were seen more like items that the user can select after her or his own taste, according to difficulty level that feels suitable. The user views a collection of video sequences, and selects those that suit her or his own needs and interests. The learner creates her own exercise. That initial solution, however, was seen as contradictory with what stakeholders wished. Because the goal of the application was to help the personnel to train measuring challenging, wide aortas, with a diameter wider than 30 mm, it was not appropriate to include easier cases. Consequently, no difficulty levels were introduced to the application. Excluding other degrees of difficulty influenced the decision about allocating functions between users and technology. As the human did not need to select how challenging exercises she wants the application became more in charge of exercises. Eventually, it was decided that the user receives a composed exercise and is supposed to examine all patients in exactly the order determined by the application.

The first solution included representing sequences as boxes marked with numbers and displaying them in a stack. Only four boxes were displayed, the topmost box represented the sequence currently played, the second box in the stack the sequence that will be played after the current one. The two remaining boxes represented the last two sequences included in the current exercise. The heuristic evaluation showed that the representation wasn't clear enough and therefore needed to be changed.

The second solution included displaying a stack of 30 boxes. Each box represented one sequence and the currently playing sequence was displayed with a contrasting background colour. The heuristic evaluation showed that readability was low and therefore the stack with 30 boxes was transformed into a wide vertical progress bar. Users who tested this third solution in form of paper prototypes could easily answer what sequence was played at the moment and how many were left.

Eventually, due to lack of vertical space, a numerical counter was introduced. Consequently the information that initially was shown graphically was presented textually (figure 2).

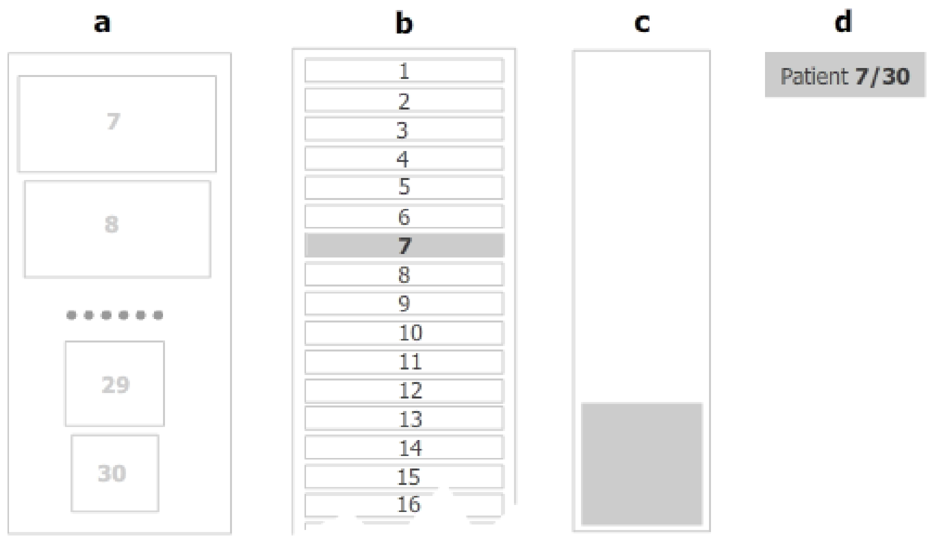


Figure 2. Displaying number sequences composing an exercise, informing what sequence is played at the moment and how many sequences are left. Four solutions were considered, see images a – d. Eventually d was implemented into the design.

Directing user toward the right image for measurement

The intention was to learn how users select the correct image to measure in from the given sequence, and then adjust the application to fit users' mental model. Observations showed that when personnel had problem with finding a clear image for measurement, they consulted a colleague and searched for a suitable image by winding a sequence back and forth. From that observed approach came the idea of directing user towards the correct image. Two ways of directing the user were proposed: directing towards colour marked areas or giving users hints to look for the correct image later or earlier in the sequence relatively to the current position of the slider (figure 3). The concept of directing the user, without providing any exact answer was liked by expert evaluators and stakeholders. Users testing paper prototypes understood the directions and acted accordingly. Eventually leading users to earlier or later parts of the sequence was chosen (B in figure 3) because it imitates the actual way to work and should therefore fit users' mental model well.

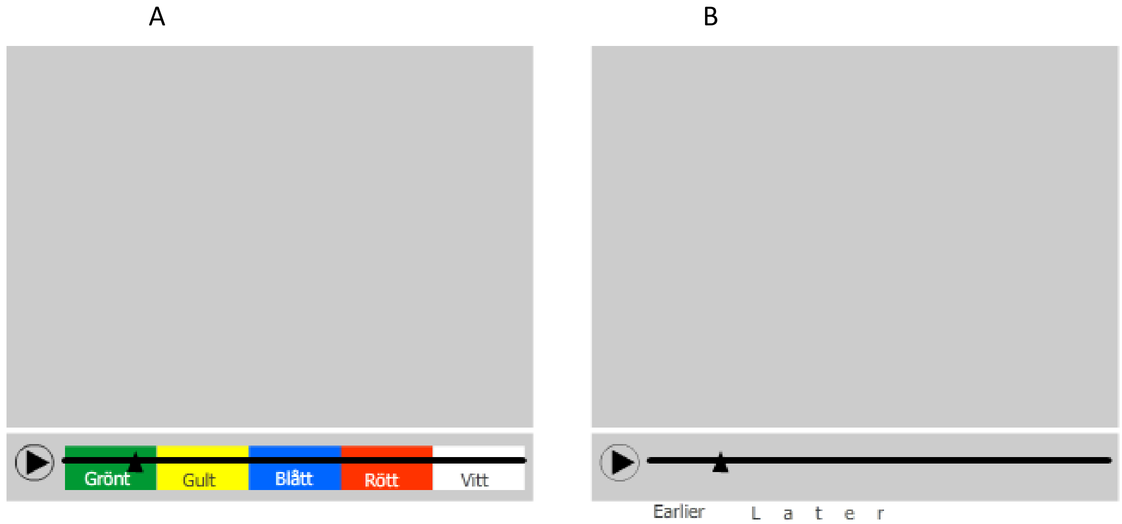
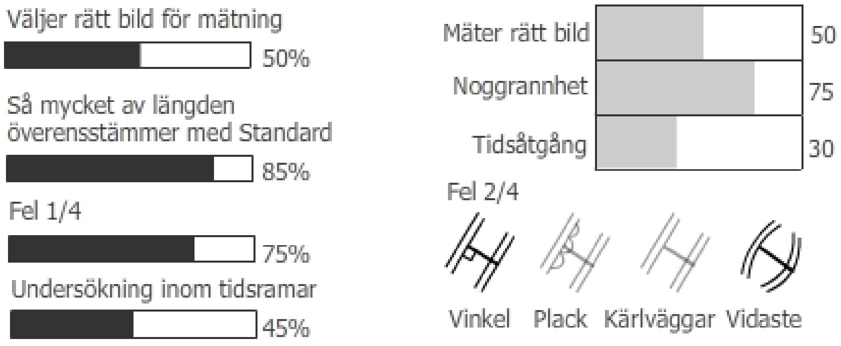


Figure 3. Two solutions to direct user toward right image for measurement:
A – directing into one of annotated, marked with colours fields, B – leading user to earlier or later part of the sequence.

Presenting user's outcome, providing understandable feedback

According to information processing theory the application helps and assists users by providing feedback and guidance. It was therefore intended to design a feedback that can easily be understood. Initially, users were rewarded with points for finding a suitable image at the first attempt, avoiding errors, accurate and quick measurement. The feedback given after each individual measurement was entirely textual but the overall result of an exercise was reported in a form of diagram.

Heuristic evaluation showed that textual feedback, given after each individual measurement, was clear but reading it was time consuming. Consequently, the textual feedback was translated into a graphical representation and represented by four and then three horizontal progress bars (figure 4-A). Bars were supposed to visually inform the user what part of totally available points was earned. In the second solution (figure 4-B) each of four errors, that the user can commit when measuring, was represented as a simple image.



A B

Figure 4. Graphical feedback. A-progress bars B-progress bars and images representing errors.

This second concept including images and progress bars was evaluated with users interacting with paper prototypes. Test participants interpreted the progress bars incorrectly. They also didn't like the idea of rewarding quick measurement as they were trained to prioritize accuracy over speed. Also the consulted senior physician mentioned that the personnel primarily should prioritize accurate measurement, without considering time.
Consequently, it was decided not to measure the examination time any more. A new solution should also include a more understandable way of presenting if the user finds the right image and how accurate the measurement was. Expert evaluators stated that despite its name (Aortaspelet, which in Swedish means “the game about aorta”) the application didn't feel like a game. Another step was therefore to make the application more game alike.

Making application more game alike

A number of digital games of different types were analysed to find a new way of providing feedback. Different types of feedback were identified. Progress bars were very common, often users were rewarded with prominently displayed points or objects for example special badges, progress was acknowledged by transferring the user to a higher level or by giving strength or better equipment. Lack of a specific object was showed by displaying an x mark beside it, a presence of an object was indicated by a tick. The results of the games review lead to a new solution (see figure 5). User's failure to find the right image for measurement is indicated by a red x mark, finding the right image was showed by displaying a green tick mark. Accuracy of measurement is shown by the prominently displayed percentage.

Figure 5. Graphical feedback – time is not measured, progress bars replaced with red x marks and green ticks.

To enhance involvement of the user a background story was created and incorporated into the game. According to the narrative the user would start the game as a nurse, medical laboratory scientist,

physician or student and aim to become a master of aorta screening who saves lives. The skills gained in the game are acknowledged by giving user a digital bronze, silver or gold badge. This story provides context and a purpose for performing tasks. A numerical counter displays the status of game completion for the player (see figure 2-D) that increases engagement. Furthermore the element of competition was introduced. Users can compare themselves with the average result within own professional group; they can also compete against the golden standard. Every examination of a patient is designed to function as a problem the player needs to solve. Outside the game, users have possibility to discuss and comment images, share hints or ask questions. It is a place where personnel with more experience can share their knowledge with beginners and students.
The idea of rewarding players with badges came from the digital game convention and the observation of personnel in the hospital. Badges are common in games, but it was observed that also personnel collected them. In computer games badges function as digital awards, they can be images of, for example a device or equipment (World of Warcraft, 2010). The healthcare personnel carries physical badges as a sign of legitimate employment status or to show their support for charity projects.

Presenting user's development

Stakeholders asked for features letting users view own development and compare with others within the same profession or within a group of people with the same experience. The first solution included presenting the results of all completed exercises in form of a diagram showing the relation between time and accuracy. User's results are displayed as a curve where the accuracy and the time of completing the game can be read. Initially four ranges of accuracy values were colour coded. Red colour denoted values between 1 and 50, orange indicated values between 51 and 80, yellow showed values between 81 and 90 and within the green area there were results with values between 91 and 100. The heuristic evaluators pointed that the most common results are placed within the upper ranges that are presented as very narrow strips and therefore the results placed within those ranges are hard to see. The four ranges of values of accuracy were therefore reduced to three, the range marked with the red colour was removed to improve readability. The result development of the logged user is displayed automatically to support the most probable use scenario, according to the Cooper's principle mentioned before. The user compares herself with other groups by applying different sets of filters. Aortaspelet's results are classified according to several attributes: profession, region, hospital and experience. Users decide which results are shown in the diagram by selecting check boxes placed on the left side of screen. By manipulating check boxes users adjust the content of the diagram dependently on actual interest and needs. Users testing paper prototypes understood the principle of filtering the results. They solved given task by selecting correct check boxes.
Even if the diagram shows how accuracy of measurements changes over time the user might also be interested in viewing details of a single game result. The percentage of cases when the user selects the correct image for measurement at the first attempt was presented together with accuracy percentage and the most common error type. These details of result of a single game were displayed in a special pane that initially was placed below the diagram and later moved over it to make it more visible.

Discussing the ultrasound images

According to the constructivist approach people learn through a discussion and cooperation with others. Consequently, the effective training application should give users possibility to exchange knowledge, questions and experiences, for instance in a forum where selected images can be commented and discussed by personnel.

3.6.2 Heuristic evaluation

Five interaction design students acted as evaluators and tested the design against ten principles for user interface design. The evaluators got a list of user interface design guidelines of Jakob Nielsen to read, then they performed tasks (see attachment) by interacting with paper prototypes. Evaluators were asked to articulate their thoughts and comment the design as they interacted with it. The test leader noted what prevented task completion, created confusion or produced error. Misinterpreting the content, not understanding navigation and assuming something is correct when it was not was noted as well. All observed issues as well as issues reported by evaluators were then categorized into low, medium and high levels of severity (see attachment). Issues directly leading to task failure were classified as high severity issues, issues contributing to task failure were rated as medium level severity problems. Issues annoying and confusing users were classified as low severity issues.

3.6.3 Paper prototype tests with the personnel

Four Medical Laboratory Scientists and two Nurses tested the proposed design. Participants were asked to perform 19 tasks (see attachment) by interacting with paper prototypes. Test participants were instructed how they should interact with the paper prototypes, for example they were supposed to simulate a mouse click by pointing an element on the sheet of paper with a pencil point. Each user was informed that it is assumed that she typed aortaspelet.se into the browser and that the first page presented is the starting point for the first task, all following tasks were started from the screen there the previous task was finished. After each action of the user, the test leader showed the paper prototype presenting the state of the system after that particular action. Test participants were encouraged to articulate their thoughts and behaviour as they interacted with the prototype. Observations and criticism given by users, problematic, uncompleted tasks and proposals for change were noted.

3.6.4 Cognitive walkthrough

The issues found during the test of the paper prototype were addressed and a new version of paper prototype was created. This new version of design was examined by a cognitive walkthrough method. A set of representative tasks was prepared as well as a list of actions needed to complete each task. The evaluator walked through all actions of user and system for each task. For each action the evaluator answered four questions presented by Dix et al. (2004). In the first step the evaluator checked if the effect of the action agreed with what user attempted to accomplish. The second question made the evaluator reflect over visibility of for example buttons and menus. In the third step the evaluator checked if the meaning and effect of a user interface element could be recognized by users. In the last question the evaluator examined if the feedback given after the action is understandable (Dix et al., 2004).

3.6.5 Usability testing

The usability testing was designed to assess how effective, efficient and satisfying the interactive prototype of Aortaspelet is. The test was prepared to evaluate if the test participants can complete given tasks, if accomplishing the tasks requires much effort and if the test participants are satisfied with the product. Planning the evaluation was based on the three usability criteria presented above, but also on user and organisational goals that were revealed during the research phase. Because the application is intended for infrequent use the intention was to optimise it for easy learning. Consequently, in addition to measurement of effectiveness (question *a* below), efficiency (*b, c*) and satisfaction (*e, f*) the learnability (*d*) was assessed as well.
The following questions were prepared to guide the planning of the evaluation and analysis of the data.

a) ***How well did the test participants complete the tasks.***
b) ***How easy it was to accomplish the tasks.***
c) ***How easy the participants thought it was to accomplish the tasks.***
d) ***How easy it was to learn the interface.***
e) ***How easy the test participants thought it was to learn the functionality of the interface.***
f) ***How satisfied the test participants were with the Aortaspelet.***

Measuring the task success required that a set of tasks was prepared. The intention was to find tasks with a clear end-state. Before the data was collected the success criteria for each task were defined. It was decided that each task has three levels of success. The success rate for each task can be: success without assistance, success with assistance and failure (the test participant gives up or wrong answer). A task was rated as successful without any assistance if the user didn't ask for help at any stage of the task. If the test subject completed the task but asked for help as she worked with it, the task was classified as successful with assistance. If the test participant gave wrong answer or gave up the task was rated as failure.

For each task a sequence of actions needed to do the task was written down. The task was considered successful if the task participant answered the question stated in the task description and/or when the last step in the predefined sequence of steps was reached.

Initially, it was planned to identify and count errors committed by the test participants. If the test person would take any step or action that prevented her from completing the task in the most efficient manner this action would be seen as an error. The identified errors would be then classified into trivial, moderate and serious. Eventually, though, when all the test sessions were completed it was decided not to count errors. As the test participants frequently asked for assistance it was assumed that many errors were prevented as they got help from the moderator. Consequently, it was decided that counting average number of times participants asked for help is more valuable. Each time a participant required assistance a special HELP marker was inserted into the session recording.

Efficiency was measured as time on task.

To measure learnability two tests with similar tasks were conducted one after another. The initial expectation was that efficiency measures like, for instance time on tasks, number of errors and success rate would stay the same in the second test. That would mean that the performance of the first test was optimal and the experience from the first test doesn't lead to better results in the second test. The differences between the performance metrics in both tests were analysed. A small gap would mean that test participants learnt quickly how to use Aortaspelet, while a wider gap would mean that learning occurs slowly.

The usability testing of the interactive prototype was performed with MORAE (TechSmith, 2010) that is a usability testing and market research software capturing the screen, recording audio and delivering surveys and tasks to test participants.

Each test participant was instructed how to manage the extra window showing task instructions displayed by MORAE. After reading the task instruction, the participants were supposed to press start task button and when they thought they were ready with tasks they should press the end task button.

The usability evaluation was an assisted testing and participants had possibility to ask for help only if they didn't know how to continue.

After each of the two usability tests, participants were asked to rate general usability of the Aortaspelet by selecting on the five point Likert scales in what degree they approve each of the ten statements of System Usability Scale (SUS). All statements of SUS were translated into Swedish. Both Learnability scale and the traditional overall satisfaction scores were counted. User satisfaction was also investigated by quick post task ratings. When test participants finished working on a specific task, they were asked to rate how difficult they thought the task was. A semantic differential scale presenting a pair of opposite adjectives was used, the first point on the scale were labelled “very difficult” and the fifth point of the scale was anchored “very easy”.

Tullis (2008) advices to assess the overall results of a usability test by comparing these results to previously established goals. Consequently, four goals, covering metrics of effectiveness, efficiency, satisfaction and learnability were established:

- Usability goal 1: all six test participants should be able to complete all tasks successfully without any help from the moderator
- Usability goal 2: no help should be required by the participants in order to accomplish given tasks
- Usability goal 3: to obtain 90 of 100 System Usability Scale (SUS) scores
- Usability goal 4: to obtain 90 of 100 Learnability scale scores

It was decided that the results of the second test were assessed as in the first testing session participants were familiarising with the test procedure including talking aloud, reading tasks and answering questions posed in the tasks, as well as completing surveys after each task.

The next chapter shows the results of the six phases of the Goal Directed Design process.

4. Results

This chapter shows the results of the Research phase, then the outcome of the Modelling phase is presented. Further, the products of the Requirements Definition phase are described, followed by the presentation of the Definition of Framework. Eventually, the results of the Refinement phase are shown, followed by the description of the final paper prototype.

4.1 Results of the Research phase

4.1.1 Stakeholders interviews

Stakeholders described the current learning style of the personnel conducting aorta screening, desired usage of the training application and needed features. Eventually they defined the user group.

Current education of the personnel conducting aorta screening

Interviewed stakeholders - two senior physicians responsible for education of personnel in the aorta screening project in the Sahlgrenska hospital pointed out that teaching of new personnel occurs in a group work. A more experienced person guides a less experienced colleague through the aorta examination and measurement. All aorta screening activities in the Sahlgrenska hospital are conducted by two employees and learning occurs through direct observation, asking questions and active participation. The existing knowledge about the abdominal aorta measurement in ultrasound images has been gathered in form of guidelines which are available for the personnel. No other learning material is used for acquiring knowledge about the screening procedure and measurement of aorta diameter in Sahlgrenska hospital.

Usage of the application

The interviews with two senior physicians indicate that the general aim of this group is a systematic, continuous quality improvement of the aorta screening program in the region Västra Götaland.
The primary expectation is that the application will make the diagnostics safer, more accurate and that the average error of measurement will be reduced. In addition the stakeholders assume that the program will coordinate and synchronize the personnel's way to work. The regular use of the application is believed to lead to more similar results between different personnel groups.
The interview persons mentioned further reasons for building the application. The program is expected to function as a continuous quality control for the experienced personnel. The regular, repetitive use of the application is believed to prevent drifting apart from actual guidelines, even in cases when a person works on her own, without possibility to consult a colleague the application will inform about current way to work.
Hospitals introducing the systematic abdominal aortic diameter examination and inexperienced personnel are expected to get quick practical introduction into the aorta diameter measurement. Another suggested idea was to use the application for the purpose of formal testing of new, beginning personnel; the test result could then determine the person's right to run aorta diameter measurement.

Functionality of the application

To satisfy needs of stakeholders the application should act as a portfolio and contain either video sequences or images of longitudinal section of aorta from one hundred patients with registered widening of aorta. Part of this collection is intended to be exchanged each year to make the exercises challenging and as actual as possible. The program is expected to be run by personnel once or twice a year; it should give a feedback on choice of image and on the measurement. Viewing own development should be possible as well as comparing with others within the same profession or within the group of people with the same experience. An introduction to aorta screening procedure and measurement for beginners was another requirement from stakeholders.
Stakeholders pointed out that to ensure transferability of knowledge gained during the training by the educational application, the behaviour of the measuring tool should be similar to the authentic

ultrasound machines. Replicating an actual work environment or elements of domain is considered important when designing a serious game and is therefore advised by Bergeron (2006).

Users of the application

The program should be available for all hospital personnel interested in aorta diameter measurement regardless the location around the country. The medical school students were pointed as the second user group.

4.1.2 Field study

User observation and user interviews

As mentioned, five types of behavioural variables were considered in the research phase. Below the results of the analysis of the interview material are shown.

Activities. The screening work starts when the personnel welcomes the person to be examined to the examination room, then man's identity card is controlled and the person is searched for in the computer register. Further, the personnel explain the examination procedure and ask the visitor to lie on the bed. The transducer is covered with an acoustic gel and placed on the stomach of the examined person. The correct segment of the aorta is identified, structures around the vessel are analysed, then walls of the aorta are recognized and electronic callipers are placed in the image. Those of the screening personnel who work in pairs consult a colleague if any problems are encountered.

Attitudes. All interview subjects had experience in aorta screening, the majority of them participated in the aorta screening project from the beginning, in 2008. All interview subjects feel confident about how to measure aortic diameter. Medical Laboratory Scientists and physicians feel more confident than nurses when handling the ultrasound machine. All interview subjects experience that obtaining a clear image is the biggest problem when measuring the aortic diameter. They also point out that it is difficult to determine walls of aorta when the patient has much fat, there is much plaque in the aorta's walls, when the aorta is snaky and when a vein or gas comes in the way. The majority of the personnel stated they are positive towards skill training by an on-line game based training application. Interview subjects emphasize that they like working together as it gives them possibility to ask a colleague whenever it is needed.
Everyone pointed out that when the aortic diameter is narrow and aneurysm can be excluded then measurement is done quickly. In contrast, when the diameter is wider more time is spent on determining walls of the aorta and plaque, in such cases a colleague is asked for advice.

Aptitudes. The direct observation of others and training under supervision were important elements of introduction into the aorta screening. When personnel want to learn more about measuring aorta they ask others or read guidelines. Some of personnel train diagnostics skills by watching ultrasound video sequences.

Motivations. Identifying aneurysm and measuring accurately is the goal of the personnel.

Skills. Medical Laboratory Scientists (MLS) have an extensive experience in using ultrasound machine. All personnel use daily the medical record system Melior (Siemens, 2010) and the patient administration system called ELVIS (ELektroniskt Vård InformationsSystem).

As mentioned earlier, the representative statements were analysed and divided into categories and ten patterns within the five behavioural variables could be identified:

a) The personnel are confident about how to measure aortic diameter.
b) The personnel are confident of handling the ultrasound machine.

c) The screening personnel work in pairs.
d) Identifying aneurysm and measuring accurately is goal of the personnel.
e) The personnel are positive towards skill training by an on-line training application.
f) The personnel are training diagnostics skills by watching ultrasound video sequences.
g) Determining edges of walls is difficult when the patient has much fat, there is much plaque in aorta's walls, when aorta is snaky or when a vein or gas comes in the way.
h) The personnel consult others when measuring (finding aorta, determining walls of vessels) is hard.
i) Obtaining a clear image (where the walls of vessels are visible) is the biggest problem when measuring the aortic diameter.
j) When the diameter is narrow and aneurysm can be excluded then the measurement is done quickly and in these cases determining where the plaque is isn't very important.

The ten behavioural variables together with the organizational and technical goals become then a foundation of the persona (see section 4.2.1).

Questionnaires

The majority, six of ten participants states that they learned measuring aorta diameter at work, the rest, that is four of ten learned measuring aorta diameter in connection with a study visit at the other hospital. If the personnel feels doubtful the measurement is repeated, the solution also involves watching aorta in both longitudinal section and cross section or watching aorta by colour Doppler. When gas obscures the aorta and makes measurement impossible at the moment the patient is asked to wait 15 minutes and do the examination later again. Ten of ten ask a colleague or a physician for advice when they feel doubtful: six of ten ask a workmate, four of ten ask a physician.
Four of ten of the personnel states there is no educational material introducing personnel into aorta measurement. Two of ten doesn't know if there is some educational material available and four of ten states there is material available at the work place.

There are guidelines (method descriptions) available in all hospitals which are one-two pages long descriptions of how the personnel should conduct the aorta examination. There is however no other material describing the aorta screening that is addressed to the personnel. Two of the interviewed stakeholders - two senior physicians stated that no educational material about aorta screening is available for the personnel except the guidelines. Consequently, the personnel acquire additional knowledge in aorta screening mainly from others. The questionnaire shows that six of ten ask others when they want to learn more about how the aorta measurement should be done. Two of ten ask others and read guidelines, two of ten only read guidelines.
All ten participants feel confident about how to proceed when measuring the aorta diameter. Physicians feel more confident than nurses. The majority, eight of ten, state that the guidance and practical training in aorta measurement, offered at work, was sufficient. According to the interviewed stakeholder, the personnel need further education within aorta examination. Half of the interviewed personnel believes they need additional education within aorta measurement. Four of ten states that no further education is necessary within aorta screening. One of ten is uncertain if any further education is needed.

4.1.3 User, organizational and technical goals

Identifying user, organisational and technical goals is essential in the process of Goal Directed design (Cooper, 2007).
According to Cooper (2007) *user goals* can be divided into *experience* goals, that is how the user wants to feel. Cooper points that, for instance no user wants to feel stupid. The second group of goals are *end* goals including what the user wants to do. The interview material shows that every interview person wants to identify aneurysm and consequently save lives. The third group of user goals are *life*

goals that define who users want to be. Based on the observation and interviews it was assumed that the personnel want to be competent at what they do.
Further, the *organizational goals* were analysed. During the interviews the stakeholders revealed that their goal is a systematic, continuous quality improvement of the aorta screening program in the Region Västra Götaland. Safer diagnostics and coordinating the way personnel work was pointed as further goals.
Making the training program a web application was a *technical goal* indicated by the stakeholders.

4.1.4 Literature review

Variability of measurement by ultrasound

The aim of the study performed by Dijnér (2010) was to examine if the results obtained by the experienced personnel performing the aorta measurement differ and how much they differ. Measurement was conducted on aortas with diameter over 30 mm. Variation between the experienced medical laboratory scientist and the golden standard (a radiologist) is 5,3 %. The maximal error of measurement was 10,4 %. Consequently, the aorta aneurysm is diagnosed correctly in 95 % (Dijnér, 2010). The study confirmed that higher diameter leads to increased variability of measurements.

The definition of game

Whitton (2010) defines game as an activity incorporating all or some of ten characteristics presented below. If the aim of an activity is winning and reaching better result than others then it has an element of *competition*. Games should be *challenging* that is offering tasks that are difficult and that require effort from the player. *Exploration* of the simulated environment and discovering constraints of the game and the interface is another characteristics of a game. The location of the game, the character and the story can include elements of *fantasy*. Additionally games usually have explicitly formulated *goals* presenting to the users purpose of the activity. Games are *interactive*, that is user affects the state of the game and gets back a response. The *outcome* of the player is the progress towards the game's objective, for example scores can be used to show user's achievements. Furthermore games require participation of other *people* in the activity, too. A game has *rules* that guide and restrain the player's actions. Eventually, the results achieved in a game have no consequences in the real world, games are therefore *safe* environments.

Engaging digital games

Whitton (2010) describes also elements influencing users’ engagement with games. The factor of *completion* makes the player wants to achieve all tasks and earn a reward. To support completion an overview of available tasks should be presented to the player as well as a reward. An engaging game has an element of *competition,* giving users possibility to compare themselves against others. The competition should however be applied carefully to avoid discouraging some learners (Whitton, 2010). A *narrative* provides a background and a context for the game. The story presents a purpose for performing tasks. Also the element of *puzzle-solving* increases engagement of the player. A *community* space in form of a forum can motivate some learners. Users can communicate outside the game to, for example share hints or ask questions and consequently get encouraged to continue even if the tasks are challenging (Whitton, 2010).

Serious games and their characteristics

Bergeron (2006) defines a serious game as an engaging interactive application with challenging goals and scoring system, that also aims to teach the player a skill, knowledge or attitude that can be used in the real world.

Several issues need to be considered when designing serious games (Bergeron, 2006).

Subtle details within the domain, like images and environments need to be replicated as accurately as possible in a game, therefore designing a successful and useful serious game requires access to the domain knowledge (Bergeron, 2006). Sometimes domain and task specific software and hardware need to be integrated to engage the learner (Bergeron, 2006). Designing a serious game requires analysis of the current learning style of user group. The most essential part of serious games is an assessment offering stakeholders possibility to evaluate whether their investment in serious games results in the desired behaviour change (Bergeron, 2006).

Examples of serious games

Serious games replicate the reality to different extent; the level of fidelity can range between low and high. Some serious games use a domain specific hardware, for example mannequins, other attempt to mimic authentic environments.

Body Simulation is a simulation based desktop application where an anaesthesiologist treats a virtual patient. Body Simulation is used both by experienced anaesthesiologists to review their knowledge and by students as a learning tool to analyse different scenarios. The user can administer fluids, drugs in different scenarios and see how the human cardio-pulmonary system responds, an operating room environment contributes to the richer learning experience (Advanced Simulation Corporation, 2004).

The user of *Cardiac Arrest!* plays a role of emergency physician. The goal is to perform resuscitation and stabilize shock of 45 patients. The player can see the patient's medical history and ECG on the computer screen, the user needs to interpret the vital signs, laboratory data and decide what treatment should be applied. The player needs to treat 15 easier patients, 15 complex patients, and 15 pediatric patients, several learning modes are available (Mad Scientist Software, 2010).

eDose teaches and assesses medication dosage calculation skills. This web application is used by nursing students to learn drug dosage calculation and by healthcare practitioners to assess dosage calculation competency. Users of eDose can see an immediate result of their dosage and evaluate their skills against the standards, the program helps to understand the basic principles related to medication orders, administration and methodology. It also helps to determine what knowledge the learner is lacking (METI, 2010).

Ultrasound imaging system

Ultrasound imaging systems are widely used due to low cost in comparison to other systems, portability and lack of known risks to the patient (Prince & Links, 2006). The primary component of an ultrasound imaging system is the *transducer* which is responsible for both generation and reception of sound pulses (Prince & Links, 2006). Prince and Links (2006) explain that the acoustic gel couples transducer to the body and enables the acoustic wave to spread into the body. The wave hits surfaces which can either scatter or reflect the sound and as the acoustic waves return to the transducer, they become converted to an electrical signal (Prince & Links, 2006). This electrical impulse is then stored, amplified or displayed on the screen (Prince & Links, 2006).

4.1.5 Examination of competitive interfaces

Picsara is an image management program used in the Region Västra Götaland. The application gives users possibility to store and analyse images, it also provides tools to measure length, area and angle in the images (Euromed Networks, 2010). After placing electronic callipers in the image, the user receives an immediate feedback in form of result displayed next to the measurement, in form of text labels. Adjusting the end points and direction of the measuring line as well as measuring radius are possible.

4.2 Results of the Modelling phase

4.2.1 Persona

As mentioned, (section 3.1.1) ten patterns within the five types of behavioural variables were identified in the research phase. In the modelling phase the findings of the research phase were synthesized and models were built.

Interview subjects were mapped to each of ten identified behavioural variables and placed on the axis of behavioural variables. However, only one behaviour pattern could be identified. No major differences between personnel groups were identified. Nurses stated that they have less experience in handling the ultrasound machine, but at the same time they confirm that they feel confident about measuring aortic diameter. Two interview subjects (Medical Laboratory Scientists) were less positive toward using an on-line skill training application, two interviewed MLS worked on their own, even if they had possibility to consult each other as they worked in a neighbouring examination room. Also, two interviewed physicians worked on their own but they had possibility to consult a senior physician. Other behavioural variables were similar for all interview-persons and therefore only one persona was created.

In the next step the goals of persona were identified: to identify aneurysm and measure accurately, to obtain a clear ultrasound image quickly, to have a colleague available to ask questions and resonate together about difficult questions, to become better at measuring.

The ten behavioural variables identified during the research phase are described in greater detail below. When differences between interview subjects were identified a behavioural axis is shown.
The figures represent different professional groups:

Figure 6. The three professional groups: Medical laboratory scientist, nurse and physician

Personnel is confident of how to measure aortic diameter

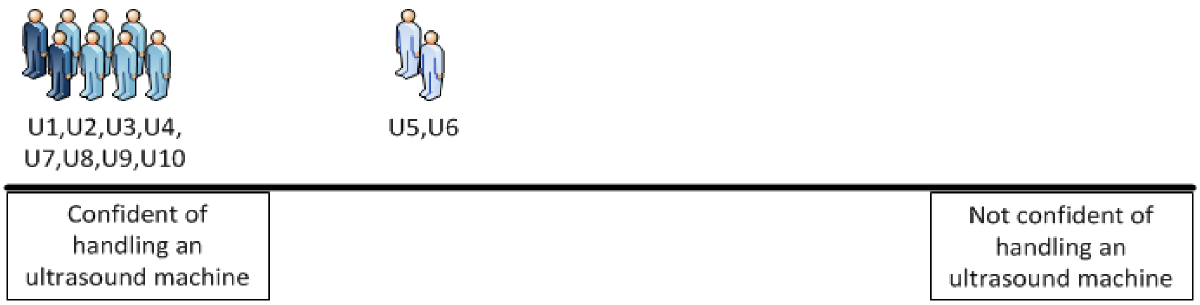
The interview material shows that the personnel feel confident of how to measure aortic diameter. Screening personnel, with a medical laboratory scientist background, relies on the own experience of performing examination of vessels. Nurses pointed that they lack the extensive experience of vessel examination but at the same time they stated that they use all available techniques, offered by the ultrasound machine functionality, to ensure themselves that the examination is conducted correctly, for example pulse Doppler and comparing diameter of cross section and longitudinal section. When uncertain how to measure, they consult a clinical laboratory scientist or a physician. Interviewed physicians rely on their general knowledge about vessel diagnostics.
All interview persons point that to know where to place electronic callipers and measure right it is essential to watch the vessel as it moves and to analyse walls and all adjacent structures carefully.
Below the most representative quotes from interviews conducted in five hospitals are presented.

Original quote in Swedish	**English translation**
”har kört mycket kärl innan”	”I have run many examinations of vessels before”
” Man får känslan för det när man tittar på det i rörelse”	“You get a feeling for it when you watch the vessel as it moves”
” får man en fin aorta så är det lätt att mäta”	“It is easy to measure if aorta is nice”

” man ser själv så att man känner att man har sett tillräckligt”	“You know when you feel that you have seen enough of the image”
”De bitar man har för att bekräfta sig själv är att tvärsnittet och längdsnittet är rimliga att det är några millimetrar som skiljer, många gånger är de lika och sen pulsdoppler”	“The pieces we have to confirm our measurement is that cross section and longitudinal section are reasonable, that there is few millimetre difference between them,many time they are similar and then pulse Doppler”
”man ser att aortan pulserar lite, att man har kollat så att det inte är cava som går och trycker ihop då vet man har sin aorta, och att man har ett pulsativt flöde när man ser i färg. ”	“ you can see that aorta is pulsing, that you checked that cava isn't pressing, then you know that you have your aorta and that you have a pulsing flow when you're watching in colour”
”du får lita på dig själv, på ditt eget omdöme, man gör ju hur man har lärt sig”	“ you have to trust yourself, your judgement, you do as you've learnt to do”
”när jag känner att jag knappt ser aortan så frågar jag någon annan”	“ when I feel that I can hardly see the aorta I ask someone else”

Personnel is confident of handling an ultrasound machine

Personnel with a clinical laboratory scientist background have extensive experience of vessel diagnostics, often they work daily with the ultrasound machine. Nurses, however perform examinations with ultrasound infrequently. Interviewed physicians handle the ultrasound machine regularly when they are on duty, sometimes, shorter periods of time they work exclusively with ultrasound examinations.

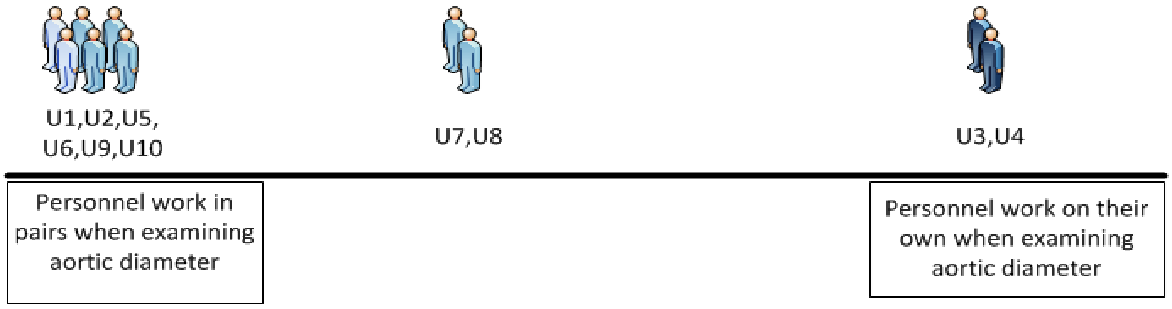


Quotes illustrating this difference in experience of handling ultrasound machines are presented below.

Original quote in Swedish	English translation
“ vi kör ju vanligtvis kärl, tromboser, hals och ben så man har ju en vana av hur kärlet brukar se ut så”	“we usually examine vessels, thrombosis, throat and legs and we have experience of how the vessel tend to look like”
”den här maskinen har vi haft på Kärl dagligen”	“this machine we use for vessels every day”
”Detta är inte apparatur som vi använder dagligen” ”vi tittar inte dagligen på ultraljudsbilder utan vi gör detta när vi screenar aortor, vi gör det en eftermiddag i veckan”	“This is not equipment we use every day” “ we don't watch ultrasound images daily, we do that only when we screen aortas, we do that one afternoon per week”

Screening personnel work in pairs

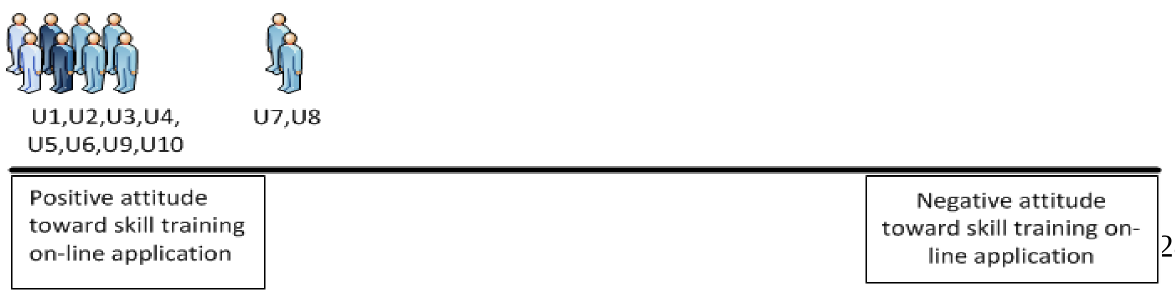
Observations and interviews show that all interviewed personnel conduct screening together with a colleague, in pairs, all emphasize that it is important to be able to consult problematic and troublesome cases with others. Sometimes both screening performers stayed in the same examination room and took turns in examining men. In one hospital, though personnel stayed in adjacent rooms and worked on their own. Despite working separately the two persons still had possibility to consult each other about problematic cases. Consequently, even this organisation of work was classified as working together, in pairs. In contrast, the observed physicians are assisted by nurses who do not measure the aortic diameter but welcome and register the patient. The figure below shows how working in pairs, when examining the aortic diameter, is distributed among the observed personnel.



Identifying aneurysm and measuring accurately is the goal of the personnel
Identifying aneurysm and measuring precisely is a goal of all interview persons.
Below, one quote from each hospital is presented.

Original quote in Swedish	English translation
”Hitta de som är i riskzonen (...), få så verkligt resultat som möjligt”	“To find those who are in risk zone (…), get as true result as possible”
”vårt mål är att mäta rätt och inte råka i alla dessa fällor. Det skulle vara synd att missa någon som är i kritisk eller i gränsfall för de kan utveckla något.”	“our goal is to measure right and not fall into all these traps. It would be pity to overlook someone who is a critical or doubtful case beause they could develop something.”
”göra så bra mätning som möjligt”	“do as good measurement as possible”
”att få en tillförlitlig mätning, hitta de som har vidgad aorta”	“to get a reliable measurement, find those who have widened aorta”
”att identifiera om det finns en aneurysm”	“to identify if there is a aneurysm”

Personnel is positive toward skill training by using an on-line training application
Majority of the interviewed persons have been positive toward a web based training application. Two interview persons were concerned that a web based solution could replace the existing training and active participation of a novice and a dialogue with others.



Quotes illustrating the positive attitude toward a web application are presented below.

Original quote in Swedish	English translation
”det här hade varit perfekt att öva.”	“that would be perfect to practise in this way”
”en lösning där man själv kan jobba skulle vara perfekt”	“a solution there you can work by yourself would be perfect.”
”Vi är positiva”	“We are positive”

Training diagnostics skills by watching ultrasound video sequences
Five of the interviewed persons train vessel diagnostics skills by watching, analysing and describing video sequences or images of vessels. The training material is provided by the External quality assurance in laboratory medicine in Sweden (Equalis) or available on the hospital's intranet.

Original quote in Swedish	English translation
”(...) filmerna vi har på intranätet, du måste gå och kolla på våran SAR webben. Där finns det filmer om åderbråck, artärer, blodtrycksmätningar och allt möjligt. De är väldigt användbara”	“Films we have on our intranet, you must see SAR webben. There are movies om rupture, arteries, blood pressure determination and many other things. Those are very useful.”
”(...) equalis.se (...) som har utbildning i allt möjligt så just för oss på kärldiagnostiken.”	“(...) equalis.se (…) that has education about many things for us working with vessel diagnostics.”
”(...) Equalis och då gör de det på kärl, man testar sig”	“Equalis and they do that for vessel, you test yourself.”

Determining edges of walls is difficult when there is much plaque in aorta's walls, when aorta is snaky, when a vein or gas comes in the way or when the patient is obese.

Original quote in Swedish	English translation
” man ska vara säker att det är plack så att det inte är skugga, det är väldigt lurigt man kan få venen som kliver in i samma bild”	“you should be confident that this is plaque and not a shade, this is very tricky you can get the vein inside the image”
” det är svår att få en bra bild pga annat som ligger i vägen, andra strukturer, artärer, vener, organ, tarmar oftast”	“ it's hard to get a good image because of all structures, arteries, veins, organs and mostly intestines”
” ibland kan venen vara väldigt dominerande så att den skymmer artären också. Det här är typiskt svårt fall”	“ sometimes the vein can dominate and hide the artery. This is a typical difficult case”
”svårt ibland att avgränsa var är vägg och var är tätstruktur precis ovanför som inte tillhör aortan, det är absolut, det är det svåraste egentligen”	“Sometimes it can be difficult to define where is the wall and where is the solid structure that is above the aorta and that doesn't belong to aorta. This is the absolutely most difficult”
” det svåraste är när de är feta, då är det långt till aortan”	“ the most difficult is when they are fat, then there is far to the aorta”

The personnel consult others when measuring (finding aorta, determining walls of vessels) is hard.
Every interview person consults a colleague or a physician when obtaining a clear image or identifying walls of aorta is impossible. Below quotes confirming that assumption are shown:

Original quote in Swedish	**English translation**
” hittar vi ingen aorta så har vi riktlinjer att vi ska kontakta kärlläkare så gör de CT”	“ if we don't find any aorta then our guidelines say that we contact a physician (?) who does CT”
” när jag känner att jag knappt ser aortan så frågar jag någon annan”	“ when I feel that I hardly see the aorta I ask someone else”
” vid tveksamma fall tar vi hellre en dialog, då har vi en intresserad kärlläkare som tar ställning”	“ in doubtful cases we prefer to consult a physician (?) who decides ”

Obtaining a clear image (where the walls of vessels are visible) is the biggest problem when measuring the aortic diameter.
Interview persons point finding a clear image as the biggest problem with measuring the aortic diameter.

Original quote in Swedish	**English translation**
” det svåraste när vi mäter är det är bildkvalitet, hur lätt kan du se kärlet. Det kan vara bilden, maskininställningarna eller patienten”	“the most difficult when we measure is image quality, how easily you can see the vessel. It can be the image, settings of the ultrasound machine or the patient”
” det svåraste är när bilderna är dåliga”	“ the most difficult is when the images are bad”

When the diameter is narrow and aneurysm can be excluded then measurement is done quickly
Quotes illustrating this approach are shown below

Original quote in Swedish	**English translation**
”om kärlet är friskt, är det 15 eller 16 så är noggrannheten oviktig”	“it the vessel is healthy, it is 15 or 16 then accuracy isn't important ”
”har man 16, 17 så gör man snabbt, man kan inte sitta en kvart med det”	“if you have 16 or 17 then you do quickly, you can't sit a quarter with it””
” bland normalvärdena, är det 17, 18 mm har man god marginal till gränser så gör det inte mycket, så resonerar vi, eftersom det är screeningverksamhet”	“Among normal values is it 17, 18 mm then you have good margin, we reason in this way, because it's a screening”

Eventually, interviewees were asked what functionality the application should have. They wanted the system to give feedback concerning the angle of measurement. Additionally, they suggested that possibility to move the callipers that were placed in the image is essential. Two interviewees pointed out that they require a program that does not require extensive mouse handling, as they had pain in their wrists.
Attitudes toward the application were mostly positive; some interview persons were concerned that a web based solution would reduce the possibility to consult others, that the training program could replace the existing training including active participation and a dialogue with others.

A third-person narrative

Identified characteristics (behavioural variables) and user goals composed a basis for the a textual portrait. The persona describes Erica, a 45 years old Medical Laboratory Scientist (MLS). Work setting is presented together with what contacts and relations with others she has at work. The user description shows Erica's goals and competence as well a her attitude toward further skill training (see attachment).

4.2.2 Workflow model

The created workflow model (see attachment) shows that measuring aortic diameter starts with finding the right segment of the aorta, if the correct section was identified then the walls of the vessels are sought. Further, the examination performer checks for plaque in the walls of vessels, analyses curves, shades and other visible structures around the vessel. Eventually, the widest part of the vessel is identified and callipers are placed in the image. The result of the measurement is documented and if the diameter is higher than 30 mm then the Sahlgrenska University Hospital is informed. If any of the steps appears to be problematic a colleague or a physician is consulted. In troublesome cases the physician may decide that the patient should be examined on another date or with another method, for example by computed tomography or magnetic resonance.

4.3 Results of the Requirements Definition phase

4.3.1 Problem and vision statements

The problem and vision statements describe user and organisational goals in a formal way (Cooper, 2007).

The problem statement defines the problem that is experienced by the persona and organization. The following problem statement was created:

Measuring abdominal aortic diameters that are wider than 30 mm is difficult and measurement results vary between the personnel in the Sahlgrenska University Hospital. If the screening personnel had possibility to train their skills that would help them to meet their goal of identifying aneurysmal aortas.

The vision statement determines the high level design purpose.

The new application will help users achieve more accurate measurement results by giving them possibility to train measuring wide aortic diameters in a safe environment which will improve the quality of diagnostics and lead to more similar measurement results between the personnel and fewer events of death caused by the rupture of the AAA.

4.3.2 Persona's expectations

The second step of the Requirements Definition phase involved summarising expectations of the persona. The following expectations of the persona were identified. The persona Erica is an experienced user of the ultrasound machine who expects that the measuring aortic diameter on the new application will be similar to the way it is done on the ultrasound machine, for instance lifting and moving callipers should be possible. It was observed that when the walls of the aorta were difficult to identify the callipers were lifted and placed in another position, personnel could relocate the callipers several times before they decided to save the image. When the quality of the image is low, for example due to the long distance between the surface of the transducer and the aorta or when there is much plaque in the walls of the vessel, the persona expects she can search for the appropriate image to measure in by rewinding the sequence back and forth. When measuring wide aortas, the persona experiences that it is difficult to maintain the right angle between the wall of vessel and the measuring line, she expects the new application to provide feedback on the angle of measurement. The ultrasound machine confirms that the selected image is saved in the data base by playing a short sound, the persona waits for the sound and sees it as an important feedback.

4.3.3 Context scenarios

When expectations of the persona were identified a context scenario was built upon user and organizational goals.
The context scenarios included the following requirements.

- Train measuring aorta diameter in ultrasound images through a web application, on any computer.
- Watch video sequences on the computer screen.
- Stop a video sequence on a desired frame.
- Get a feedback on the choice of the frame.
- Measure the diameter in the chosen video frame.
- Place callipers in a chosen image and see the diameter length in millimetres displayed.
- Get from the program an immediate feedback on the placement of callipers.
- See what mistake(s) were made when placing callipers.
- Get to know how to avoid mistakes in the future.
- Read existing rules and guidelines about measurement of aorta diameter.

The context scenario describes in very general terms how the persona interacts with the training application as she measures the aortic diameter. The scenario contains no details about the interface, just its general behaviour and how it satisfies the persona's needs.
The text below shows that the context scenarios show interaction in general terms and on a high level.

> *"When Erica feels she has enough information about aorta and adjacent structures she stops the sequence at the most suitable image and starts measuring. She places the callipers and can see the length in millimetres displayed. She gets an immediate feedback on the placement of callipers. Beside her callipers the correct callipers are shown in a distinctive colour. She sees clearly that she didn't place callipers in the right position. The program informs Erica what mistakes she did and how she can avoid them in the future."*

4.4 Results of Framework Definition phase

4.4.1 Form and posture

The stakeholders decided that the application should have a form of a Web application, that should be viewed on a high resolution computer screen.

4.4.2 Data and functional elements

Three basic elements were identified: video sequences, exercises and results (see table 1). Each exercise contains 30 video sequences that are named with numbers. It should be clear if a sequence was accessed before or if it was recently added to the collection. Exercises should be marked, e.g. with numbers, the user should know what exercises were accessed. One possibility to satisfy different groups of users could be to provide several levels of training: one for beginner, e.g. students, intermediate and more challenging advanced). Users need an overview of when each result was obtained, if several results are displayed it should be clear who owns what results. Viewing the best result of the owner is desirable as well as seeing the best result within a user group. The most recent result should be displayed in a distinguishing way, automatically.

Objects -basic units	Video sequences	Exercises	Results
Attributes			
Attribute 1	Video sequence number	Exercise number	Date of the result
Attribute 2	Accessed before yes/no	Difficulty level (beginner, intermediate, advanced)	Owner of the result
Attribute 3	Added recently to the collection yes/no	Accessed before yes/no	The best result of the owner
Attribute 4			The best result within a user group
Attribute 5			The most recent result

Table 1. Data elements

Afterwards, possible operations on data elements were analysed (see table 2). The user should be able to play, stop and pause a video sequence. The number of completed video sequences should be displayed for the user. Users with no or very little experience of aorta diameter measurement would maybe need easier exercises to stay engaged and interested in training therefore the idea of providing different difficulty levels. The application should inform what exercises are completed and when they were completed. Starting and ending an exercise is another operation that is necessary. Operations on results include viewing the result of just completed single video sequence as well as viewing the results of a completed exercise. The user should be able to view and publish the best result. The summary of all completed exercises should be available as well. Users should be able to view an average result within a group of users.

Functional elements	Video sequences	Exercises	Results
Operations on data elements			
Operation 1	Sorting sequences according to difficulty level (beginner, intermediate, advanced).	Sorting sessions according to difficulty level (beginner, intermediate, advanced).	Viewing the result of a just completed single video sequence.
Operation 2	Displaying video sequences completed before.	Displaying all completed exercises, providing exercise name/number and date of completion.	Viewing the result of just completed exercise.
Operation 3	Starting and playing a sequence.	Starting an exercise.	Saving an exercise result.
Operation 4	Ending and pausing a sequence.	Ending an exercise.	Publishing the best result.
Operation 5			Viewing the best exercise session result.
Operation 6			Viewing an average session result within a group.
Operation 7			Viewing the results of your all completed sessions.

Table 2. Functional Elements

4.4.3 Functional groups and hierarchy

In this step, the identified data and functional elements were grouped into tasks. The possible operations were listed in chronological order, arranged into screens (see table 3), described in a key-path scenario and then, in the next step, presented visually as mockups.

In this step it was decided that the training application will have only one difficulty level, therefore the operation of choosing a difficulty level of an exercise wasn't represented as a mockup in the next step.

Screen	Operation
1	Creating a user account
2	Starting a new exercise
3	Choosing a difficulty level of an exercise
	Getting an overview of video sequences composing the chosen exercise
4	Starting a video sequence
	Pausing a video sequence when the right image appears
	Accessing measuring tools: electronic callipers
	Placing callipers on the image
	Getting the length in millimetre displayed
	Seeing a response callipers displayed
	Seeing the length of diameter compared to the length of gold standard, see what mistakes were made and the error percentage after measuring diameter in just completed sequence
	Starting the next sequence
	Saving the result of an exercise or auto saving
5	Seeing an overall result of just completed exercise.
	Viewing results of all completed exercises.
	Comparing own results with an average result within a group.
	Publishing the result of an exercise.

Table 3. Chronological order of operations grouped into screens

4.4.4 Key path scenarios

Key path scenarios describe how the persona manipulates elements of the interface to achieve her goals and in what order. Key path scenario describes how the persona performs eight tasks and shows in detail all actions of the persona and the system. The tasks described by the key path scenario include creating a user account, logging in, starting a new exercise, choosing a frame from the video sequence, placing electronic callipers in the image, ending exercise, viewing summary of overall results and getting to know the scoring system. Below a description from the key path scenario is shown. The interaction of persona Erica with the training application is described as a task of choosing a frame from the video sequence is performed.

> “Choosing a frame from the video sequence.
>
> Erica views the sequence and decides to stop it at frame number 30; she believes that the chosen frame is the most appropriate for measurement. As she pushes the pause button

on the video player panel, the feedback panel displays a text message informing her that she chose to measure on a frame placed in the green area and that the correct frame is placed in the yellow area. Erica manipulates the slider of the video player, goes to the recommended yellow area and views that section of the movie. As she stops the video sequence again the feedback panel displays a text message informing that the chosen frame was the right frame to measure on and that she can start placing electronic callipers within the image."

4.4.5 Validation scenarios

The following list of infrequent user actions were created:

- the user wants to end the training application before the game is completed
- the user asks the system to validate the measurement before the correct frame is selected
- the user wishes to view result summary as she is playing the game
- the user wants to change the experience in own profile
- the user wishes to change the hospital in the own profile
- the user selects incorrect frame 20 times

All infrequent scenarios were addressed even if no mock-ups were created to demonstrate them.

4.4.6 Visual Design

The logo is a combination of a circle and a concave line. The red and white colour palette was applied. The intense red colour can be associated with the bleeding that is a consequence of a ruptured aortic aneurysm. The concave line represents the aorta, the circle stands for an aneurysm, and the white dotted line in the middle of the circle represents the measurement line placed by the screening personnel.

The *Adjacency* colour scheme provided by Microsoft Visio was selected for the prototype.

4.5 Results of the Refinement phase

In this section the results of testing mock-ups for usability is shown. First the outcome of heuristic evaluation is described, followed by the results of testing paper prototypes with end users. Further, the effects of the cognitive walkthrough are shown. Finally, findings of the formal usability testing of the interactive prototype conducted with the end users are described.

4.5.1 The results of heuristic evaluation

As it was mentioned before, the collected usability issues were categorized into low, medium and high levels of severity issues.
Issues that directly lead to task failure were classified as *high* severity issues. The most serious problems were related to navigation. Evaluators couldn't for example find where to actually train measurement and where to read about how to measure. Misleading terminology was another group of issues having impact on effectiveness, for example use of the terms *game* and *exercise* as synonyms confused evaluators and contributed to errors.
Several issues were connected to displaying results of a single game. Evaluators didn't understand how many points they were given and what errors they made, consequently the task couldn't be completed.
Issues contributing to task failure were categorised as *medium* severity issues, for example viewing points earned in the game that was played for some time ago was problematic due to small text. As a consequence, completing tasks took longer time and was annoying.
Issues frustrating and annoying users were classified as *low* severity issues, for instance evaluators didn't like wording used in the instruction and asked for a bigger play button in the media player panel.

As second and third design versions were created the number of high severity issues declined (see figure 7 below). Five of eight issues identified in the third version of design were classified as low severity issues. The low number of high severity issues in the third version of design was seen as a sign of improvement.

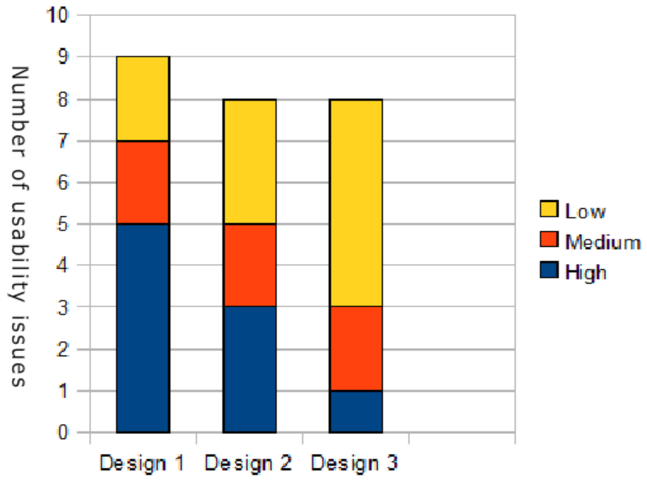


Figure 7. Number of unique usability issues by design iteration. Unique issues observed by heuristic evaluators. Severity ratings low, medium and high show how serious the issue is.

Heuristic evaluation showed that evaluators didn't understand why the term game was used as a synonym to exercise; they didn't see the programme as a game even if users were rewarded with points. This finding showed that it is important to work on making the application more game like.

4.5.2 The results of testing paper prototypes with end-users

Testing paper prototypes with users revealed seven usability issues. Below, all issues are described together with solutions applied.

1. The user didn't understand when a new game was started. Even when a new game began the participant couldn't tell that the task was completed and attempted to press “start a new game” button. Placing a bold title showing the game number directly under the main menu is believed to make it clear that a game's started.
2. When asked to leave the game and view the results the participants were confused if to press the “save game” or the “end game” button. One possible way to solve that issue is to provide only one “end game” button and make the programme save the progress automatically.
3. Another issue was related to the progress bar visualising finding the right image for measurement. Test participants couldn't interpret the progress bar correctly. An alternative way of showing if the user found the right image for measurement involves using the tick and x-marks (see figure 5).
4. When test participants failed finding the right image for measurement they were instructed to try again. It was observed that they didn't know how to proceed; eventually they found the solution and decided to press the play button. A textual hint added into the instruction explains how to proceed and saves user's time.
5. The participants felt that assessing their own experience could be difficult; classifying themselves into one of four experience groups was subjective and unclear. They suggested that reporting their own experience could be done according to number of examinations done by ultrasound during the last year. The same suggestion was given by one of the stakeholders.

6. Test participants were confused by a reward clock displayed in the corner. It was interpreted as an actual clock indicating how long time the game was played. Consequently a new reward was created in form of a bronze, silver and gold badge.
7. Participants noticed also that measuring the time of the game is unnecessary since to measure right is more important than to measure quickly. The same observation was done by a stakeholder; consequently the time reporting was removed.

Testing paper prototypes with future users revealed several positive usability findings. These observations were helpful when the new design was created; they influenced the decision about what elements should remain unchanged.
Participants commented that a progress bar is growing when they moved to the examination of the next patient which indicates that a progress bar was noticed and interpreted correctly. Participants understood that a click on the circle of the curve, in the result section, is needed to display the results of the previous game. Participants commented that if they were measuring very well they would publish their own results, which indicates that they understood the function of the “Show my results to others” button. Participants followed written instructions and understood how to proceed when the application directed them to different sections of the colour coded slider. When asked what error was committed participants answered correctly which indicates that the presentation of committed errors was clear. Progress bars showing time and accuracy were understood correctly as well, even if the time measurement was seen as unnecessary. Participants committed no errors related to top level navigation and when filtering results.

4.5.3 The results of cognitive walkthrough

It was found that when users created a new account the starting page for the game was displayed and no confirmation was given telling that the creation of a new account was successful. Consequently it was decided that a confirmation message should be displayed for users.

Analysis of actions needed to comment an ultrasound image showed that commenting an image required that users typed in their own name. This action was seen as unnecessary and users’ own name should be displayed automatically, however, at the same time users should have possibility to change this default value and remain anonymous.

Examination of steps needed to view the average accuracy of measurement of experienced medical laboratory scientists showed that no results were displayed to the user until all three parameters were selected: profession, hospital and experience. If user selected only a profession no results were shown on the diagram. This lack of direct feedback from the system can be confusing and therefore it was decided that default values should be displayed, for example experience and hospital matching user's own parameters.

When a game was completed the user clicked the “show results” button that displayed the result section. This system action was considered not explicit enough. Completing a game is an achievement and should be rewarded accordingly. Therefore a separate modal panel displaying accuracy and reward was created and presented to the user on the foreground of the results section.

The findings of the cognitive walkthrough lead to a new design that is described in the section 4.6.

4.5.4 The results of the usability testing

Below the questions defined in the method chapter are answered.

How well did the test participants complete the tasks

There is a difference between the task success in the first and the second test sessions. In the first test, the participants required assistance to complete the tasks, in the second testing session, the test subjects could complete more tasks without asking for help.

In the first test, tasks three, four and five weren't completed by all participants (see figure 8) - one participant failed to complete the task three, one participant failed to complete the task four and two participants failed to complete the fifth task. Three participants completed the first, second and fourth task without any assistance from the moderator. The third and fifth tasks were completed without any help by two participants. The remaining test subjects needed help to complete the tasks successfully.

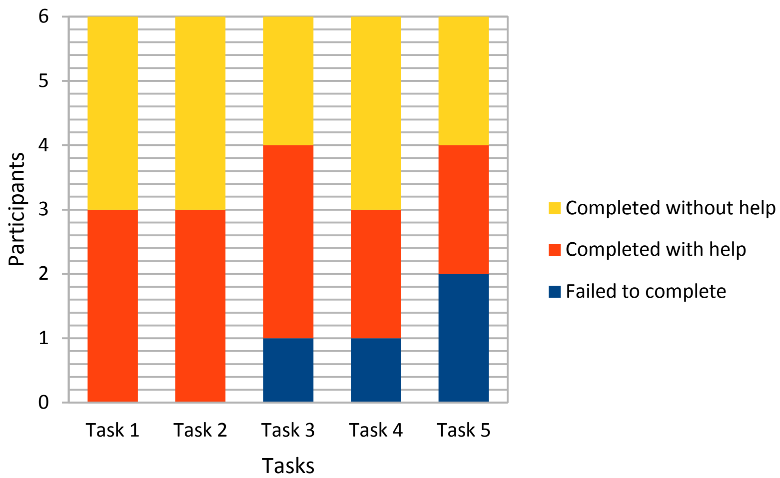


Figure 8. Success distribution by task in the **first test**.

In the second test, one participant failed to complete the fifth task (see figure 9). One participant completed the fourth task with help from the moderator. Two test subjects needed assistance to

complete the second task. Assistance was provided to the two participants working on the third task. All six participants completed the first task without any help. Task two was completed successfully without any help by four participants, ratings for task three are the same. Five participants accomplished the fourth task with no help. Similarly, five participants completed the fifth task successfully, without requiring any help.

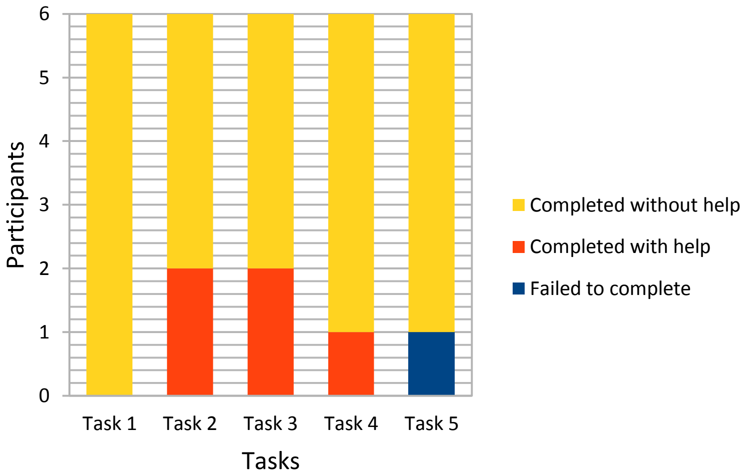


Figure 9. Success distribution by task, in the **second test**.

How easy it was to accomplish the tasks

Participants often asked for help. During the first test, the test subjects asked for help more frequently than during the second test (see figure 10). During the first test, on average, 1,5 questions were asked during the third task, in the second test this number occasions when the participant needed help declined to 0,3. During the second test, none of the test participants asked for help as they worked on tasks one and five. In the second test, the test subjects asked for help on the second task almost as often as in the first test (see figure 10).

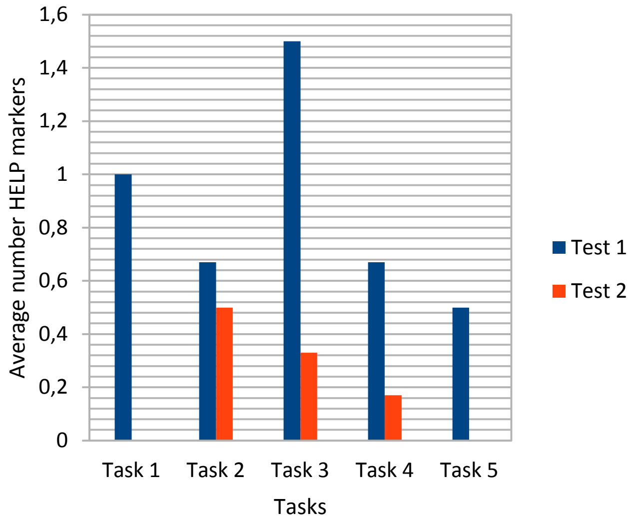


Figure 10. Average number times the participants needed help by task

How easy the participants thought it was to accomplish the tasks

When test participants finished working on a specific task, they were asked to rate on a 5 points scale how difficult they thought the task was. Selecting 1 meant “very difficult” and selecting 5 indicated “very easy”. The average scores given by the participants during the second test were higher which means that the participants thought the tasks were easier then (see Figure 11). Only the first task was experienced as more difficult in the second test.

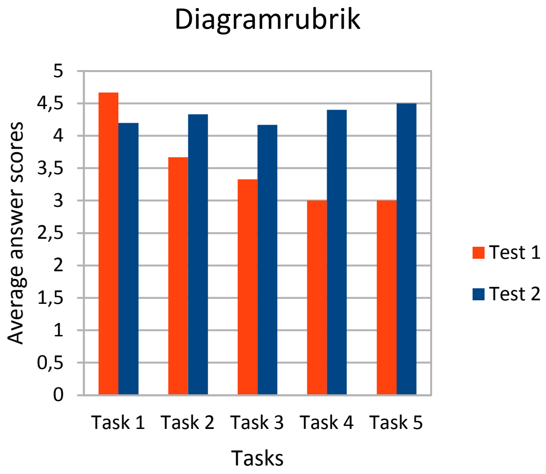


Figure 11. Average answer scores by task in the **first** and **second** test.

How easy it was to learn the interface

Time that participants needed to work on tasks was shorter in the second testing session (see the figure 12). During the second test, the test subjects asked for help more rarely than during the first test (see figure 10). In the second testing session, more test subjects could complete tasks without asking for help (see figure 13). For tasks one and five the performance of test participants improved greatly, consequently, the gaps between the task success rates for the first and second test are relatively big (see figure 13). In contrast, the gap between the task success rate for task two, for the first and second test is the smallest (see figure 13).

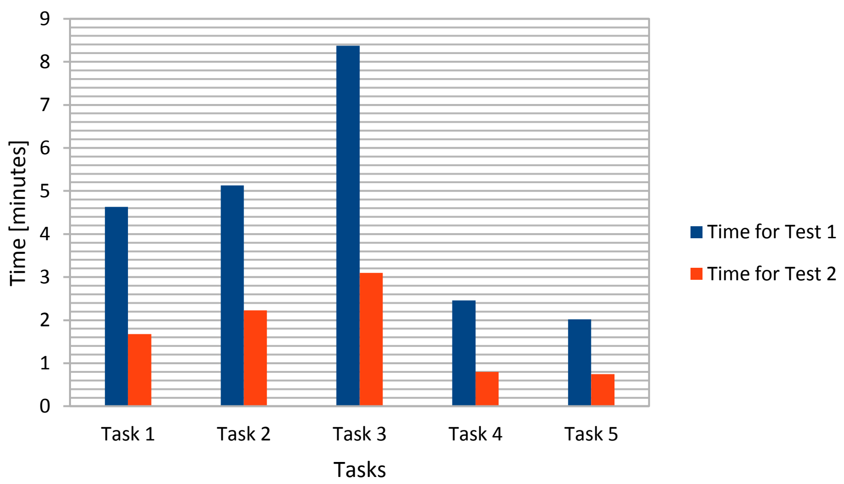


Figure 12. Average time per task (minutes) during the first (Test 1) and the second testing session (Test 2).

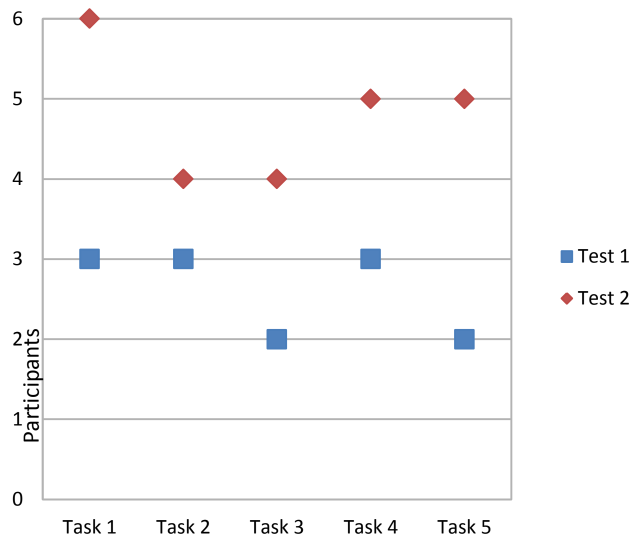


Figure 13. Task success without assistance by task, during the **first** (Test 1) and **second** (Test 2) test session.

How easy the test participants thought it was to learn the functionality of the interface (post test rates).

In the first test the participants rated the learnability of the Aortaspelet as 60,4 on the 100 points scale. After the second testing session, the test subject rated the learnability scores to 58,3. After the first test, the perceived learnability score of the application was better than after the second usability test.

How satisfied the test participants were with the Aortaspelet

The average of the SUS scores given by the test participants, after the first usability test, was 55 on the 100 points scale. After the second test, on average 58,7 points were given.

The SUS scores given by the test participants were higher in the second test (see figure 14). Consequently, the personnel testing the Aortaspelet perceived the application more usable after the second time they interacted with it. Only one test subject (U6) gave lower usability scores after the second testing session.

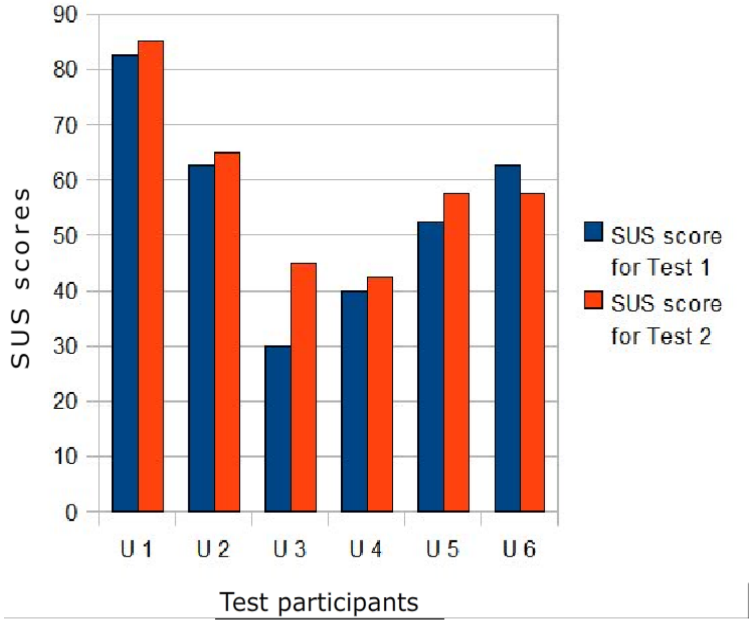


Figure 14. SUS score given by the six test participants (U1 – U6). The scores were collected after the first (Test 1) and after the second test session (Test 2).

Questions asked by the test participants revealed usability issues

When working on the first task, that is when the participants created a new account and logged in, the test subjects were missing supportive text clearly explaining what kind of input was expected from them.

As the test participants worked on the second task, that is when they explored the information on the web site to find out what are the requirements for getting a gold badge and what the third principle of measurement, the main issue was concerning the phrasing of the question, additional explanation and rephrasing was required.

The third task, that is playing the game, revealed issues with play, pause and stop buttons. Any of these controls was labelled and some test participants didn't recognise the functionality of these buttons. As the ultrasound sequence was played very quickly, the participants didn't reach the pause button before the sequence was half way; consequently, viewing the sequence was done primarily by dragging the slider. Going forth and back through the sequence required steady and smooth hand movement and watching the slider as it was manipulated, participants commented that both watching the slider and the screen was challenging. One participant suggested using a mouse scroll wheel for viewing the sequence. When test participants placed the callipers some of them wished to view the length immediately. They explained that before they save the measurement, they measure in several places of the image to determine where the vessel is widest. Repositioning callipers was challenging for some participants. The tooltips explaining how to handle callipers didn't appear quickly and participants attempted to move the callipers in the same ways they do on ultrasound machines (by clicking a calliper once and then dragging it). Further, active buttons “Show results”, “Control if the image is right for measurement” and “Control my measurement” weren't noticeable enough, but the meaning of the buttons was understood.

The fourth task, that is viewing details of their own result, revealed issues concerning filtering and viewing the results on the diagram. The text explaining that the circle should be pressed wasn't visible or understandable. It wasn't clear that clicking a circle displays the results in the panel.

During the fifth task the participants were asked to find out how accurate the experienced medical laboratory scientists tend to measure. Some participant working on this last task didn't understand that items selected by manipulating the check boxes were displayed in the diagram. In some cases the circles displayed on the diagram just weren't noticed and understood.

The table below presents unique usability issues revealed by the questions asked by users. Questions of test participants were translated into English.

Task	Help was concerning	Issues that were revealed
Task 1	Do I need to type my name and surname or can I be anonymous?	The fields that require inserting data weren't defined and marked.
	How much experience do I have?	Users weren't supported by any data when they assessed own experience.
	Should I choose my own password here?	There was no text informing users that the password should be created by them.
	What is my username, I don't remember.	Text explaining that an email address works as a username wasn't placed above the username text field.
	Why can't I log in? (When logging in the test participant typed username she uses at work into the text field.)	Text explaining that an email address works as a username wasn't placed above the username text field.
Task 2	I can't see the answer here. What should I do now?	The question wasn't clearly phrased.
	Should I find this out through playing a game or should I read some information?	The question was ambiguous and needed to be explained and rephrased.
	Where can I find information about the rules of the game?	It wasn't clear that the information about the rules of the game could be found in the Introduction section.
Task 3	Should I close this page before I open the game page?	Participant's mental model included a rule that one page needs to be closed before a new page can be opened.
	Can I make the sequence go slower?	The sequence was played to quickly.
	Do I stop the sequence with this button?	Play, pause, stop buttons weren't labelled.
	It is impossible to scroll the sequence back, isn't it?	Handling the slider was difficult and jerky.
	Why can't I view the sequence? (Test participant tried to drag the stop button instead of the slider)	The slider wasn't visible enough or it didn't appear to be draggable. The stop button didn't convey its function in a clear way as it wasn't labelled.
	Can I scroll the mouse scroll wheel to view the sequence?	Test participants are used to using special scroll ball placed in the keyboard of the ultrasound machine, the mouse scroll wheel was suggested as an alternative to dragging the slider.
	Can I change the placement of callipers after I placed them?	The tooltip describing how to handle the callipers took too long time to appear, the test participant didn't notice it.

Task	Help was concerning	Issues that were revealed
	Can I see how much my measurement is?	The participant explained that she wants to measure in several places to decide where the vessel is widest, then she wants to save the measurement she knows is the widest. The length of the measurement is not displayed immediately after the callipers are placed in the image.
	What happens with the callipers?	In her daily work the test participant places one calliper in the image and then drag the cursor to the second edge of the vessel, when the button was released the second calliper was placed. The tooltip took time to appear and didn't inform user that each calliper is placed in the image by a single click.
	Where is the measure button?	The test participants are used to the interface of the ultrasound machine where after finding the right image for measurement a “measure” button is pressed to activate the measuring tool (callipers). Test subject acted as if it was a typical screening for aneurysm and not a training game with feedback functionality. The purpose of the Aortaspelet wasn't conveyed in a clear way.
	How do I save the image now?	When examining patients for aneurysm, the test participants measure the aneurysm and then save the selected image into the data base. It wasn't clear enough how the test participant should continue after placing callipers in the image.
	To continue should I press “New game”?	The “View result” button, letting users to continue the game wasn't noticeable enough.
	Can I measure the same sequence again? (after I got my feedback)	It was unclear how the user should continue.
Task 4	I can see where the circle representing my results is but I can't see what error I did.	The text explaining that the circle should be pressed wasn't visible or understandable.
	What does it mean *“details”* of my results?	Term “details” is too general.
	When did this appear?	Test participant was concentrated on the diagram and didn't see that the results were shown in the panel above the diagram. It wasn't clear that clicking a circle displays the results in the panel.
Task 5	Where do I see the result when I select the check box?	It was unclear that the diagram contains and displays the items that the user selects by manipulating the check boxes. Circles weren't noticed.
	I am an experienced BMA, do you want to know my results?	The question was ambiguous and needed to be rephrased, that the question was to find out how medical laboratory scientists as a group measure.
	Should I insert my data into check	It wasn't clear that by selecting check boxes the data

Task	Help was concerning	Issues that were revealed
	boxes? I already did that when I created my account.	is filtered and not inserted.

Table 4. Unique questions asked by users and usability issues revealed shown for each task.

4.6 The final paper prototype

This section shows the final design in form of a paper prototype.

4.6.1 Starting page

On the starting page users can see a video clip presenting the application, test playing the game, create a user account, log into their own account or read about the game project.

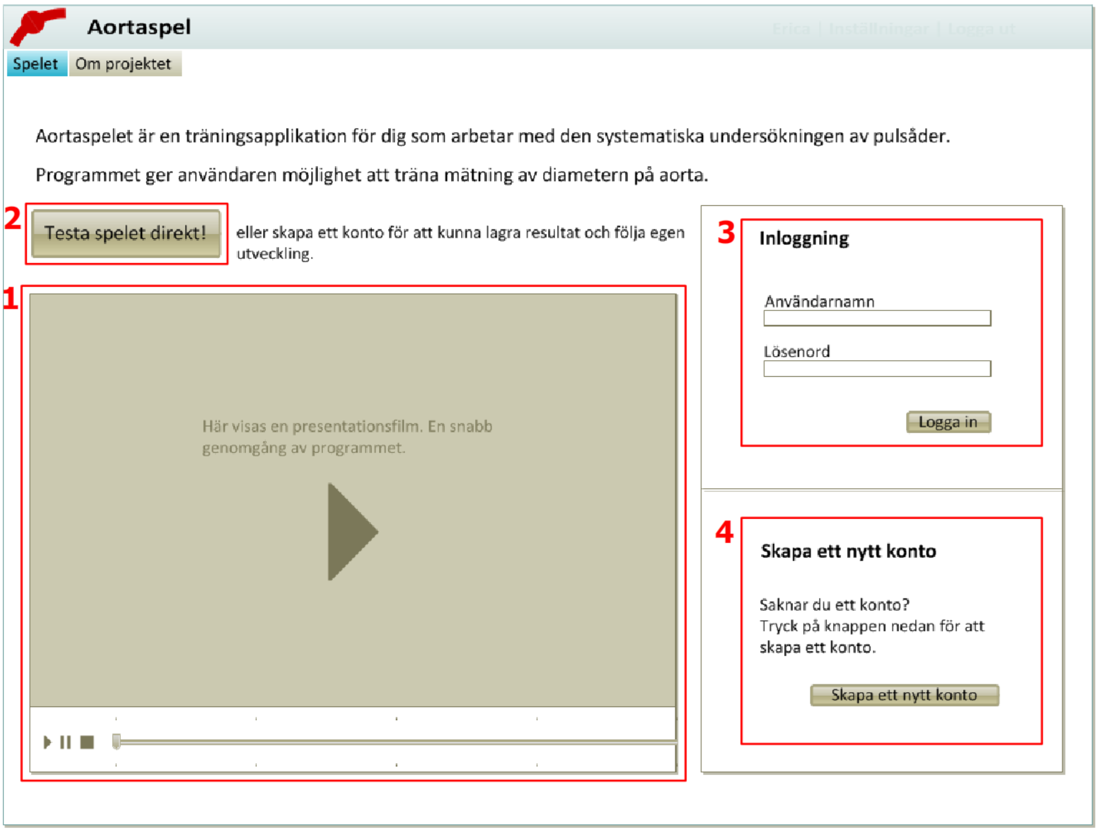


Figure 15. Starting page: 1-video clip presenting the game, 2-button lets user test the game without logging in, 3-log in text fields, 4-user can create a new account.

Areas 3 and 4 shown in figure 15 are arranged asymmetrically, visually prominent elements, that is titles and buttons, are placed in upper-left and lower-right corners. This arrangement creates visual flow that makes the user's eyes move from the top-left to the bottom-right that suits users using left-to-right languages (Tidwell, 2006).

To ensure that users complete the “logg in” form quickly eye movements need to be reduced. To quicken the form processing time the input fields and corresponding labels (see area 3 in figure 15) were vertically aligned (Wroblewski, 2005).

4.6.2 Creating a user account

A wizard leads users as a new account is created. A task of creating a new account was chunked into three simple steps where users fill out forms collecting data about personal information, experience and username and password. Wizard consists of three groups of elements (see figure 16): 1- a sequence map at the top shows the user an overview of all steps and where he or she is in the steps, 2- input text fields, 3- navigation back and forth. Elements creating the wizard are arranged symmetrically. To help people connect each label with its associated control the form design follows the right-left alignment pattern, that is the labels on the left are aligned to the right and the elements on the right are left aligned (Tidwell, 2006, Wroblewski, 2005). Additionally the right-left alignment saves the vertical space and makes that all the content of the form is displayed above the page fold which helps users to avoid scrolling (Nielsen, 2010, Wroblewski, 2005).

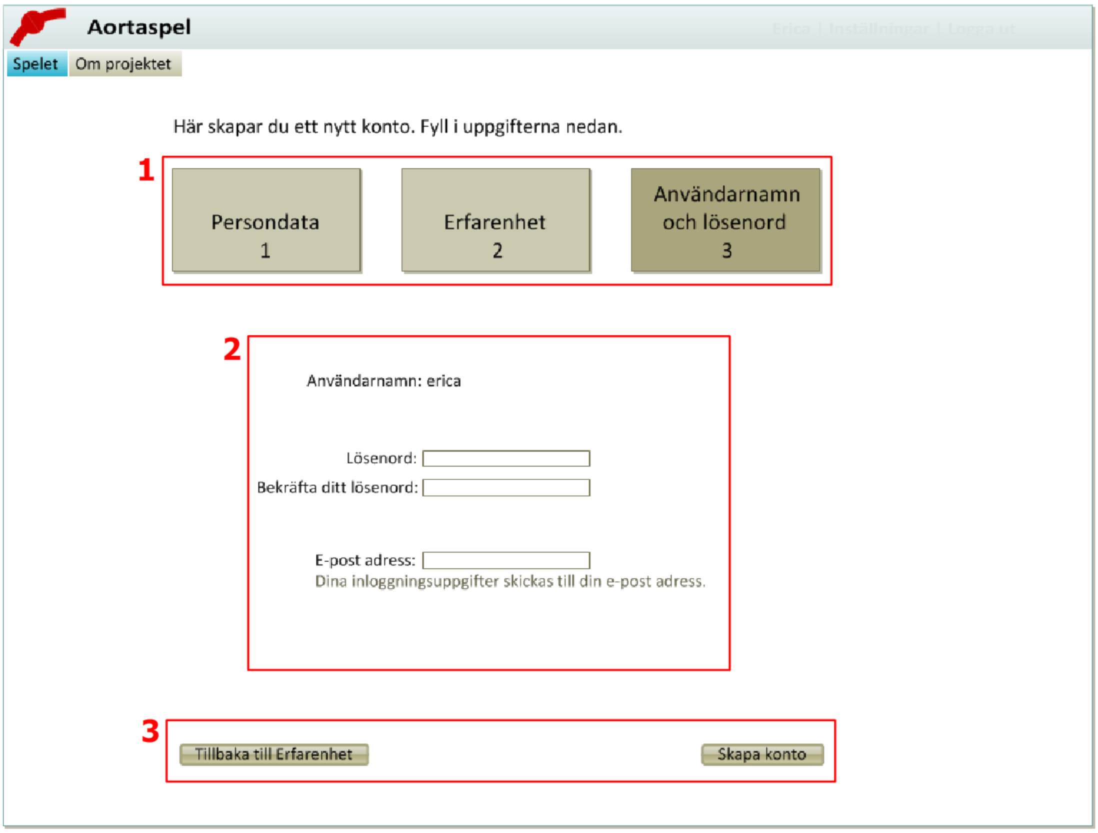


Figure 16. Wizard helping users to create a new account. 1- overview of all steps, 2- input text fields, 3- navigation back and forth.

Navigation

Primary navigation allows users to navigate quickly among main groups of content, global navigation is consistently displayed on all pages and therefore always visible and accessible for the user (Vora, 2009, Tidwell, 2006). The global navigation in Aortaspelet is placed horizontally at the top of the page and divides information into four main categories: *introduction, play, results* and *discuss images*, user can access these groups from anywhere within the application (see 1 in figure 17).

Secondary navigation demonstrates other navigation options within the selected primary navigation category (Vora, 2009). Secondary navigation in Aortaspelet is placed vertically and runs down the page on the left side.

In the Aortaspelet the global navigation is always placed at the top of page and local navigation on the left side of the page, this repetitive and consistent inverted-L navigation layout contributes to logical coherence, harmony and uniformity through the application.
Utility navigation provides two key functions that always can be accessed: logout and user settings. The name of the owner of the account is displayed in the utility navigation bar as well (see 3 in figure 17).

Introduction (introduktion)

The *introduction* category of the global navigation contains only three subcategories: rules of the game, measurement principles and information about goals of the project (see 2 in figure 10). Entries to all those groups could be placed in a horizontal local menu, below the global navigation bar, however to keep consistency and logical coherence through the Aortaspelet the left side of the page was used for the local menu.

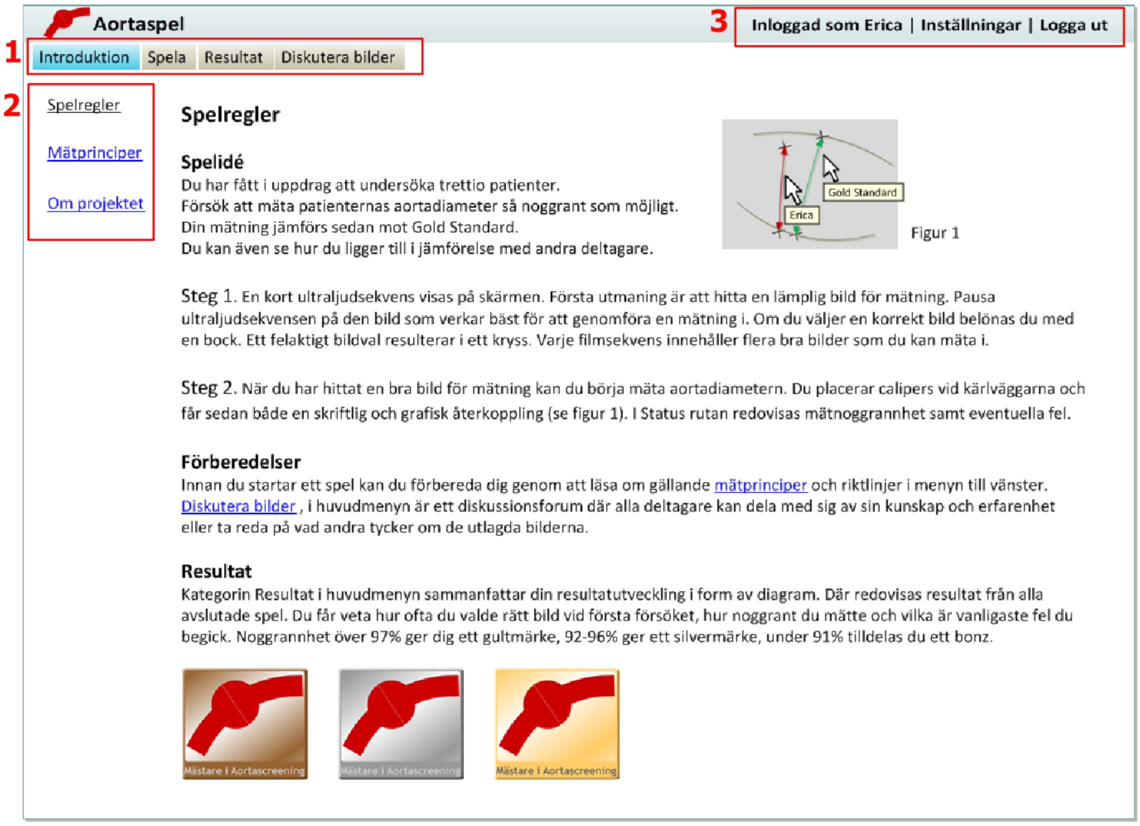


Figure 17. Introduction page and navigation, 1- primary navigation, 2- secondary navigation, 3- utility navigation.

4.6.3 Play (Spela)

The second global navigation category, the *play* section, uses the left side of the page to display critical information for gameplaying such as textual instructions for the user and player's outcome (see 2 and 3 in figure 18). User's failure to find the right image for measurement is indicated by a red x mark, finding the right image is showed by displaying a green tick mark. Accuracy of measurement is displayed in a bold style and error type is shown as an image.
The central part of the page is devoted to the video player screen; user can control the video sequence by manipulating a slider and pressing pause, play or stop buttons.

Directly over the video player screen there is a group of two buttons: “Is this image suitable for measuring” and “Control my measurement” (nr 4 in figure 18). A panel placed beside the button group shows how many patients were examined, the next patient can be accessed by pressing “Next patient” button (nr 5 in figure 18). The user can end the game, start a new game or access an uncompleted game by choosing options from the toolbar (nr 6 in figure 18).
A prominent title informing user what game is being played is displayed directly under the primary navigation bar.

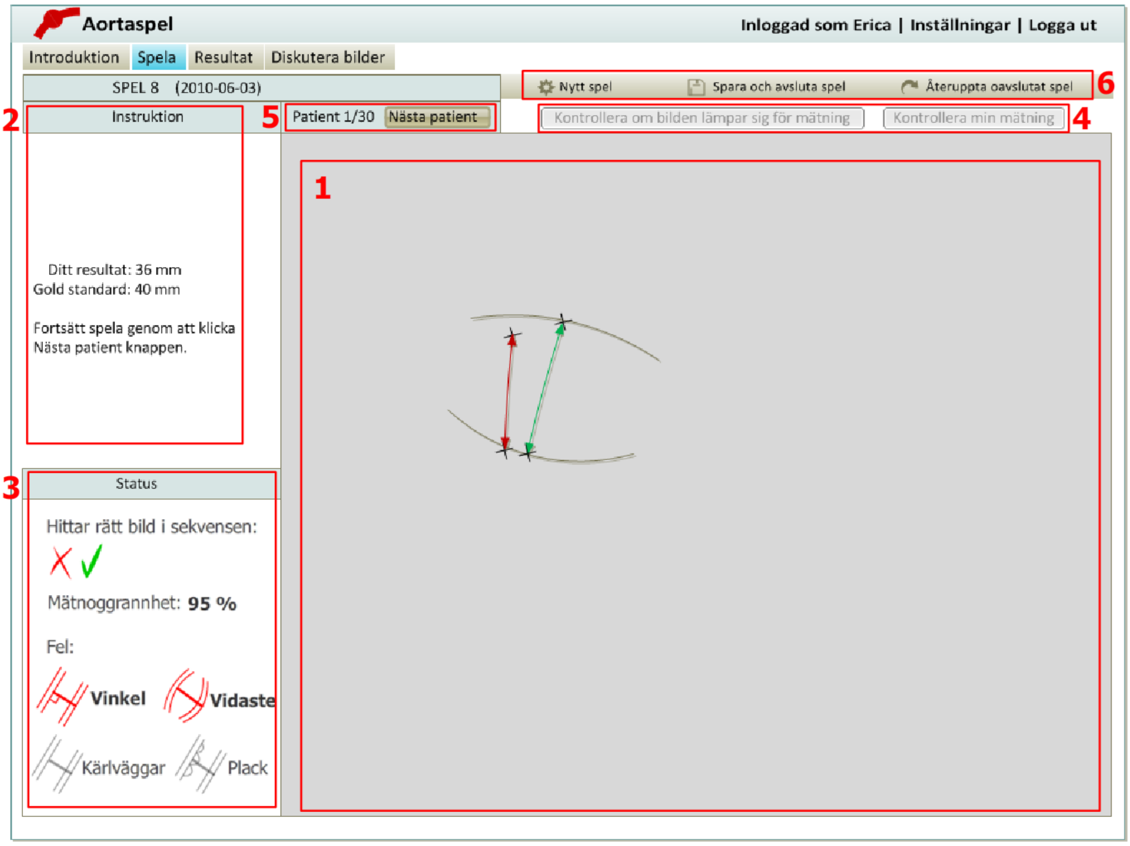


Figure 18. The play section, 1- the video player screen, 2- an instruction and feedback for the player, 3- user's outcome, 4- a group of two buttons “Is this image suitable for measuring” and “Control my measurement”, 5- shows how many patients were examined, the next patient is accessed by pressing “Next patient” button, 6- toolbar letting user manipulate manipulate the game.

4.6.4 Results section (Resultat)

The third global navigation option is labelled *results* and displays player's outcome (see figure 19). It also informs about users’ progress over the time and provides means for comparison between results of different groups of users. Users navigate through all the results by applying different sets of filters. Aortaspelet's results are classified according to several attributes: profession, region, hospital and experience. Users decide which results are shown in the diagram by selecting check boxes placed on the left side of the site. By manipulating check boxes users adjust the content of the diagram dependently on actual interest and needs. Users’ own accuracy curve is a dotted line labelled with the user’s name; all completed games are represented as clickable circles. When a circle is clicked additional information about the result of the selected game is shown in the panel located in the top-

right side of the page. By pressing the “Show my results to others” users reveal their best result to colleagues.
Results of others are presented as a curve displayed in a solid line style.

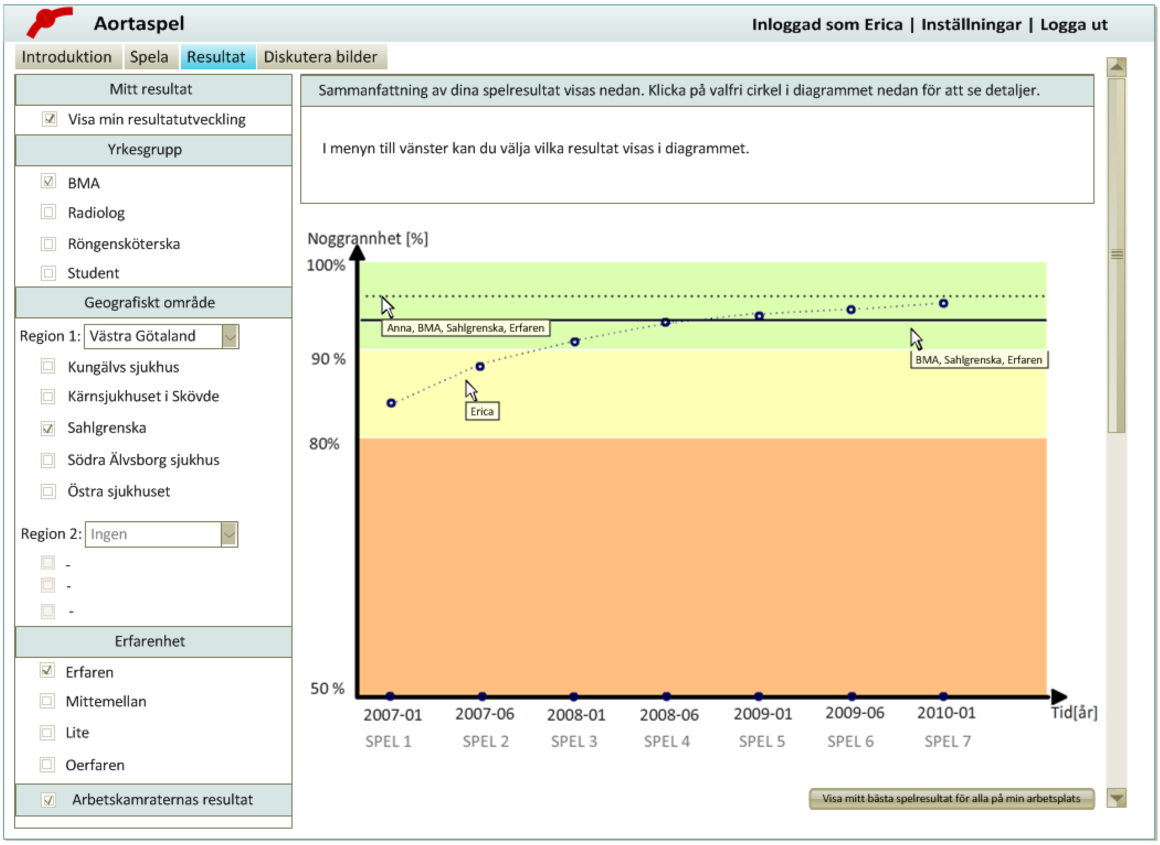


Figure 19. Comparing results of different groups of users.

4.6.5 Discuss images section (Diskutera bilder)

The fourth section of the global navigation gives user possibility to discuss images. User inserts a text and name, a comment can be also submitted anonymously (see figure 20). Every image is presented together with a number of comments written by a radiologist and users.

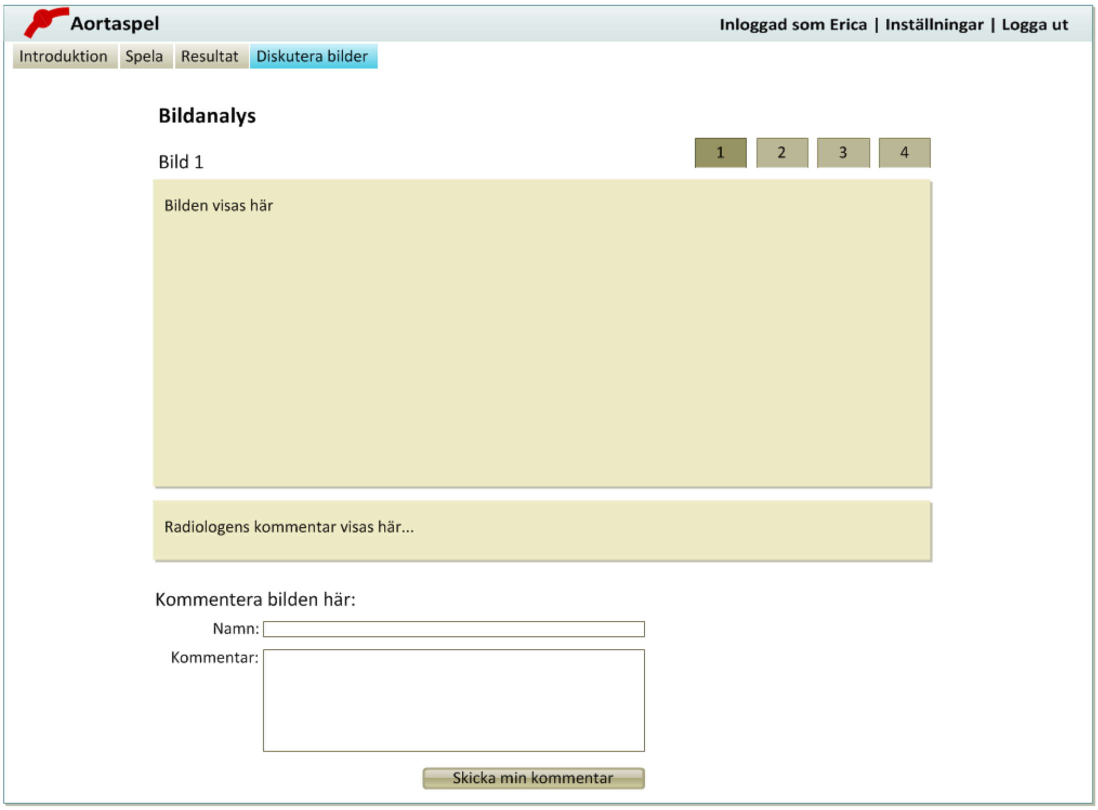


Figure 20. Discussing and commenting images.

4.7 The interactive prototype

The interactive prototype was developed by a software engineer within one week. Although the appearance of the prototype is very alike the design conveyed by the paper prototypes, not all functions were implemented. The content of the "discuss images" section is not working. The local menu of the "introduction" section is lacking. The most essential operations, which is measuring the aortic diameter and viewing feedback, are however functioning. Viewing own results of the recently played game and seeing average results of different professional groups by selecting check boxes, in the "results" section works properly as well.

Based on the capabilities of the developed interactive prototype five tasks were prepared for usability testing with end-users (se appendix 8).

4.8 Development support phase

Paper prototypes and scenarios were presented to the developer. The developer together with the designer drew an application flow diagram as a support for programming work. When the interactive prototype was nearly ready a meeting was arranged to discuss the capabilities and limitations of the program as well as final, necessary changes.

5. Discussion

In this chapter the strengths and limitations of method application in this design process is discussed. Further, the results of the usability evaluation are assessed. Additionally, it is discussed if the Aortaspelet can be seen as a portfolio. Further, the challenges of conducting a field study and usability evaluation in the hospital environment are discussed. The chapter concludes with suggestions for further work.

Methodology
Involvement of end users was very valuable and made the design process focused on adapting the training application to the actual work practise of the screening personnel. The field study and evaluations with users directed the work and helped to prioritize.
The Requirements Definition phase appeared to be the most significant phase of the Goal Directed process; it was however based on the findings of the research phase, consequently both stages contributed heavily to the design. Especially identification of persona's expectations during the Requirements Development phase was influencing the design solution. Context scenarios and key path scenarios helped to look at the persona's potential interactions with the training application. As the key path scenarios were much more detailed than the context scenarios, they helped to picture and imagine different design solutions.

The video material that was collected during the field study helped the observer to understand the measuring process. As the examination of aortic diameter was done quickly, recording the screen of the ultrasound machine, when the measurement was done, helped to prepare questions for the interview, identify details of work and later, in the Requirements Definition phase to summarise the persona's needs.

Test of the paper prototype was valuable but it wasn't possible to test and evaluate all functions. The paper prototyping appeared to be suitable for testing if the right concepts and terminology were used, if test participants understand the navigation, content and if the functionality meets their requirements. The test based on paper prototypes appeared to be inappropriate for evaluating controls, for example if users can manipulate the slider or if they would notice information appearing in the panel after they clicked the circles on the diagram. During the task 4 of the usability test, it happened that the participants didn't notice that after clicking a circle, new information appeared in a panel. When the paper prototype was tested it was more obvious as a new sheet of paper prototype was handled to the test participant. These observations about appropriateness of paper prototyping are confirmed by Snyder (2003) who explains that the main strength of paper prototyping is that it can simulate *what* the interface can do, but it can't mimic *how* the future application will respond.

Usability testing software Morae made collection and analysis of data relatively easy and quick, one person probably wouldn't be able to conduct a usability testing including measuring time on task and number of times users needed assistance, rating task success, delivering tasks and surveys and helping the test participants whenever they needed. The software made it possible to test the design in the working place of the participants, only with a laptop and headset.

Throughout the design process a design diary in Google Wave was kept. Produced documents, collected images, encountered problems and ideas were shared with several involved persons. This common space on the web made it possible for the supervisors to give feedback and follow the process. The diary helped to review all the collected resources and eventually, to write this report as the most important material was gathered in one place.

Limitations of the method
The severity rating system is often criticised because various usability specialists rate issues differently, however in this design process the severity rating was used as a way to identify progress and to direct work to solve the most serious problems.
During the paper prototype testing with users, only one person was moderating, playing computer and making notes which influenced the quality and number of observations.

The persona represents primary users, that is the screening personnel. Stakeholders pointed students as a second group of users, however no research of this group was conducted and therefore the solution is adapted primarily for use by the experienced personnel. Another group of users that the design doesn't take into consideration are physicians who will act as Gold Standard. They need possibility to select, mark and upload the ultrasound sequences. The Gold Standard will be responsible for updating the game. No research of and solution for this user group was build.

For the needs of the actual project the field study wasn't as extensive as suggested by the contextual inquiry method. The observation lasted between one to two hours as the work with screening is only one of many activities of the personnel. No other work of the personnel was taken into consideration. The further analysis of the collected material didn't follow the approach suggested by the method, that is creating an affinity diagram. The data was instead analysed by qualitative analysis methods.

Paper prototype
User of Aortaspelet decides which results are displayed by applying different sets of filters. The downside of the filtering solution is that it is possible that for some combinations of filtering options the user will see no results at all, sometimes the answer can be an empty set (Vora, 2009), e.g. it is likely that in the Aortaspelet there are no results of users who are both “experienced” and “students”. On the other hand it is relatively easy to apply different set of filters and change filter criteria.

Usability test
As mentioned before, assessing the results of a usability test can be done by comparing these results to previously established goals (Tullis, 2008). Below, I discuss whether the four previously established usability goals were achieved.

- *Usability goal 1: all six test participants should be able to complete all tasks successfully, without any help from the moderator*

According to the goal, in the second testing session, all six test participants should complete all tasks without any help from the moderator. The target value was, however reached only for task 1 (see table 5). The second and third task were completed successfully without any assistance by four test participants. Tasks four and five were completed by five test subjects.

Task / Test 2	Actual number of test participants who completed tasks without help	Goal
1	6	6
2	4	6
3	4	6
4	5	6
5	5	6

Table 5. Actual number of test participants who completed tasks without any help compared to the goal.

- *Usability goal 2: no help should be required by the participants in order to accomplish given tasks*

Participants needed help from the moderator during the second test as they worked on tasks 2, 3 and 4, thereby the set goal wasn't met.

- *Usability goal 3: to obtain 90 of 100 System Usability Scale (SUS) scores*

After the second testing session the test participants gave 58,7 SUS scores, thereby the goal of obtaining 90 points wasn't met.

- *Usability goal 4: to obtain 90 of 100 Learnability scale scores*

After the second testing session the test participants gave 58,3 Learnability scores, consequently the goal of collecting 90 points wasn't reached.
The analysis shows that previously defined usability goals weren't met, particularly subjective satisfaction and learnability rates are far from the target values. Task completion rates were nearly reached, however the elements of the interface, handled when the tasks 2 and 3 are solved, need to be refined and improved as only four participants could complete these tasks successfully without any help from the moderator.

The combination chart below (figure 21) summarises the data from figures 9 and 11. The diagram shows task completion rate and subjective rating (0-5) for each of five tasks in the second usability test. The diagram shows that task 3 was the most problematic for participants because it has the lowest values on both scales. In task 3 test participants were asked to play the actual game, it was the most essential task that required the biggest number of steps from the user. The problems encountered by the test participants revealed issues with rewinding the sequence forth and back and repositioning callipers, both these features were required by the end users and identified in the section *persona's expectations*. The limitation of the paper prototype made it impossible to evaluate these features as they require actual handling of a slider with mouse. Even during the second test, the participants needed help to complete task 3 (see figure 10) which points that issues are serious and need to be solved.
Figure 21, below shows disparities between the task completion data and the subjective ratings, for example tasks 1 had only moderate subjective ratings but the highest task completion rates. It might suggest that even if task 1 was easily accomplished by the participants, the missing instructions informing for instance that one's e-post address functions as a username, made that participants didn't experience the task as very easy. Finally, the figure below helps to distinguish the tasks that had reasonably high values for both metrics, such as Tasks 4 and 5. This result might suggest that even if tasks 4 and 5 don't reach the goal of being completed by all six test participants with no assistance, they were seen by the participants as relatively easy to accomplish.

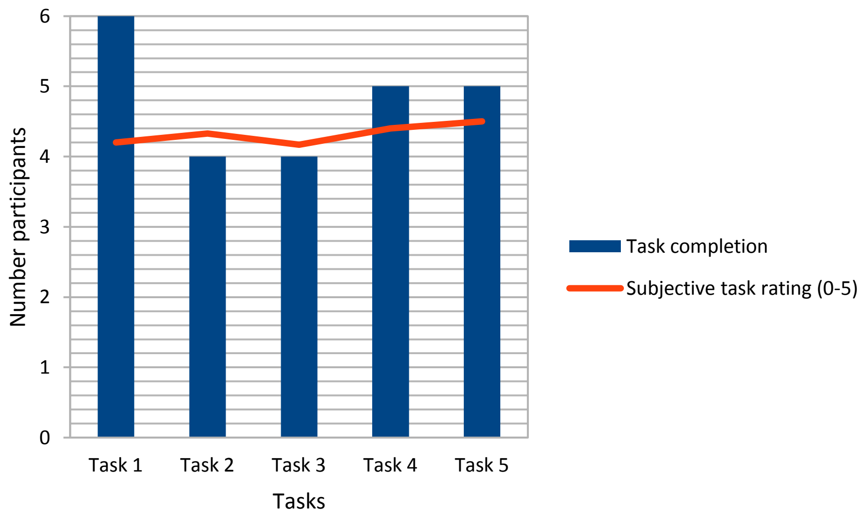


Figure 21. Combination chart showing task completion rate and subjective rating (0-5) for each of five tasks in the second usability test.

The findings of the usability testing show that the designed game based training application is not fully usable, consequently additional refinement of the interface is necessary. The usability testing lead to better understanding of interaction between users and the interface. The collected questions asked by the test participants, together with usability issues could guide the work of improving the design.

Aortaspelet as a Portfolio
Aortaspelet incorporates elements of a portfolio pedagogy. All personnel performing screening of aorta collects image material, an experienced radiologist selects representative images and describes them, every user of the application has then possibility to comment the available set of images. Consequently every image is presented together with a number of comments written by a radiologist and users.
Typically, a portfolio requires that each user personally collects, selects and reflects over artefacts and builds own portfolio that shows professional development and competence. Even if the Aortaspelet incorporates elements of portfolio pedagogy: collecting, selecting and reflecting, it doesn't give a single user possibility to influence all three elements. Therefore Aortaspelet can't be seen as a typical electronic portfolio. On the other hand it is a collection of carefully selected images that support learning and skills assessment.

Application of educational theories
All three educational theories influenced the designed application. The training program rewards desired user actions according to the behaviourist approach. Both the task of measuring aorta diameter and user's cognition are modelled and user's actions are monitored by the program which is typical for the information processing approach. Additionally, users have the possibility to discuss images within community of professionals which is an element of the constructivist approach.

A field study and evaluation in hospitals
Conducting a field study in the hospital environment required that the privacy of patients was taken into consideration. To observe the personnel in hospitals a recommendation letter was needed, it was also required that I signed a non-disclosure agreement. The visited hospitals have restrictions about usage of the hospitals' cable internet, consequently when the interactive prototype of the web application was tested a mobile broadband was used.

Future work
The online environment of Aortaspelet could be supplemented with mobile technology. An e-mail message or an sms could be sent to users to remind them that six months has gone since they played the game. In that way an online game could have an element of mixed reality game and fit user's daily routines better. Additionally, skilled users could be rewarded with a physical gold badge.

Two interview persons were concerned that Aortaspelet would reduce the possibility to consult others, consequently, when introducing the application to users the purpose and consequences of it should be clearly stated.

The experience of the personnel could be assessed automatically as the frequency of the examinations could be extracted from the hospital or regional data base.

Once the interactive prototype becomes more usable, a test should be conducted examining if the regular use of the designed application leads to more accurate measurement among the screening personnel or quicker learning of the work practise.

6. Conclusion

The goal of this master's thesis was to design a usable game based on-line training application helping the screening personnel to train measuring abdominal aortic diameter in ultrasound images.

The Goal Directed design process was applied and finally, in the end of the design process, usability of an interactive prototype was evaluated with users by studying their performance on tasks and subjective ratings.

In conclusion, the usability study conducted in the end of this design process has shown that the interactive prototype of the Aortaspelet is not yet fully usable. For instance, handling the slider and watching the image at the same time was problematic. Consequently, additional improvement of the interface is needed. The usability testing provided rich qualitative and quantitative material about the interaction between test participants and the interface. The collected questions asked by the test participants while they worked on given tasks, together with revealed usability issues could guide the next design iteration and lead to more usable design.

Acknowledgements
A number of people have generously given time, advice and information during this project. I would like to thank the people who supported me in my work:
I am thankful to my supervisors, Lars Lindsköld and Mikael Wintell from Bild- och Funktionsregistret of the Region Västra Götaland, for the guidance, feedback and encouragement during this master's thesis.
I am also grateful to Olof Torgersson, the programme director of the Interaction Design for help and advice throughout this master's thesis.
I would like to thank Göran Bergström and Johan Fredén-Lindqvist, senior physisians from Sahlgrenska University Hospital, for their feedback.
Eventually, this thesis would not have been possible without the clinical laboratory scientists, nurses and clinitians working with screening for AAA. Many thanks for explaining details of your work and participating in usability testing of paper and interactive prototypes. Thank you all for your contribution and valuable feedback.

References

Advanced Simulation Corporation, 2004. *What Is Body Simulation?* [online] Available at: http://www.advsim.com/biomedical/what_is_body_simulation.htm [Accessed on 7 September 2010]

Bergeron, B., 2006. *Developing Serious Games.* [e-book] Boston, MA, USA: Course Technolgy. Available at Chalmers Library: http://site.ebrary.com/lib/chalmers/Doc?id=10228186&ppg=38 [Accessed August 2010]

Bergqvist, D., 2007 *Bukaortaaneurysm* [online] Available at: http://www.internetmedicin.se/dyn_main.asp?page=1523 [Accessed 24 November 2010]

Beyer, H. Holtzblatt, K., 1998. *Contextual design : defining customer-centered systems*. Academic Press.

Brooke, J., 1996. *SUS: a "quick and dirty" usability scale*. In Jordan et. al. ed. 1996 *Usability Evaluation in Industry*. London: [online] Available at: http://www.usabilitynet.org/trump/documents/Suschapt.doc. [Accessed December 2010]

Brothers, S. K., 2007. *Game-based e-learning: the next level of staff training, combining the convenience of online learning with just plain fun adds a new dimension to staff training.* Long-Term Living for the continuing care professional. [online]

Available at: http://www.ltlmagazine.com [accessed 25 November 2010].

Cooper, A., Reimann R. & Cronin, D., 2007. *About Face: The essentials of Interaction Design 3rd ed*. Wiley Publishing.

Corbet, A.T., Anderson, J.R., 1989. *Feedback timing and student control in the Lisp intelligent tutoring system.* Proceedings of the Fourth International Conference on AI and education, pp.64-72.

Curtise, KC., White, P., McKay, J., 2008. *Development of web database portfolio system with PACS connectivity for undergraduate health education and continuing professional development*. Computed methods and programs in biomedicine 94 (2009), pp.26-38.

Dix, A. et al. 2004. *Human-computer interaction*. London, Prentice Hall.

Dijnér, T., 2010. *Mätvariabilitet vid screening av abdominala aortaaneurysmer med ultraljud.*

Euromed Networks, 2010. *Picsara* [online] Available at: http://www.euromed.se/html/picsara1.htm [Accessed on 6 October 2010].

Hornbaek, K. 2006. *Current practice in measuring usability: Challenges to usability studies and research.* International Journal of Human-Computer Studies 2006(64), pp.79-102.

Johansson G. & Swedenborg J., 1986. *Ruptured abdominal aortic aneurysms: a study of incidence and mortality.* British Journal of Surgery. 73 (February), pp. 101-103.

Jokela, T., 2008. *Characterizations, Requirements, and Activities of User-Centered Design—the KESSU 2.2 Model*. Human-Computer Interaction Series, 2008(2), pp. 168-196. in red. Effie Lai-Chong L., Ebba Hvannberg T., Cockton,G., 2008. *Maturing Usability Quality in Software, Interaction and Value*.

Lewis, J.,R., Sauro, J., 2009. *The Factor Structure of the System Usability Scale*. [online] Available at: http://www.measuringusability.com/papers/Lewis_Sauro_HCII2009.pdf [Accessed on December 2010].

Lindholt, J.S., 1999. The validity of ultrasonographic scanning as screening method for abdominal aortic aneurysm. *European Journal of* Vascular and *Endovascular Surgery. 17, pp.472-475.*

Mad Scientist Software, 2010. *Cardiac Arrest!* [online] Available at: http://www.madscientistsoftware.com/cadetails.html [Accessed 2 September 2010].

METI, 2010. *eDose.* [online] Available at: [http://www.meti.com/products_edose.htm] [Accessed 7 September 2010].

METI, 2010. *Human patient simulator*. [online] Available at: http://www.meti.com/products_ps_hps.htm [Accessed 7 September 2010].

Nielsen, J. 2005. *Ten usability heuristics.* [online] Available at: http://www.useit.com/papers/heuristic/heuristic_list.html [Accessed on August 2010].

Nielsen, J. 1994. *How to conduct a heuristic evaluation*. [online] Available at: http://www.useit.com/papers/heuristic/heuristic_evaluation.html [Accessed on July 2010].

Nielsen, J. 2005. *Heuristic Evaluation*. [online] Available at: http://www.useit.com/papers/heuristic/ [Accessed on July 2010].

Nivre, J., et al., *Göteborg Transcription Standard*, Göteborg: Department of Linguistics Göteborg University. [online] Available at: http://www.ling.gu.se/projekt/tal/doc/transcription_standard.html [Accessed May 2010].

Pleumeekers, H.J. et al., 1998. *Differences in observer variability of ultrasound measurements of the proximal and distal abdominal aorta*. Journal of Medical Screening 5, pp.104–108

Prince, J.L. & Links, J.M., 2006. *Medical Imaging Signals and Systems*. New Jersey: Pearson Prentice Hall.

Quintana, Ch. Et al., 2003. A framework for understanding the development of educational software. In ed. Jacko, A., Sears, A., 2003. *The Human-Computer Interaction Handbook Fundamentals, Evolving Technologies and Emerging Applications*, pp.823-834.

Sharp, H., Rogers, Y., Preece, J., 2006. *Interaction design beyond human-computer interaction*. 2nd edition. John Wiley & sons, Barcelona.

Siemens, 2010. *Melior*. [online] Available at: http://www.nwe.siemens.com/sweden/internet/se/Healthcare/IT-losningar/Melior/Pages/Melior.aspx#Generellt%20om%20Melior [Accessed on June 2010]

Siemens, 2010. *Syngo Dynamics.* [online] Available at: https://www.medical.siemens.com/webapp/wcs/stores/servlet/PSOptionProductDisplayView?catalogId=-1&catTree=100000&langId=-1&productId=178891&storeId=10001

Snyder, C. 2003. Paper prototyping : the fast and easy way to design and refine user interfaces. [e-book] Available at: http://proxy.lib.chalmers.se/login?url=http://www.books24x7.com/marc.asp?isbn=1558608702 [Accessed on August 2010]

Singhl, K., Bonaa, K. H., Solberg, S., Sorlie, D.G., Bjork, L., 1998. *Intra- and Interobserver Variability in Ultrasound Measurements of Abdominal Aortic Diameter. The Tromso Study.* Vascular and Endovascular Surgery, 15, pp.*497-504.*

Sitzmann, T., 2010. Game On? The effectiveness of game use in the workplace depends on context and design. Training & Development. Vol. 64 (10), pp.20-20 [online] Available at: http://web.ebscohost.com.proxy.lib.chalmers.se/ [accessed 25 November 2010]

Totty, M,. 2005. Better learning through gaming. *The Wall Street Journal Online* [online]. Available at: http://www.qube.com/pdf/whitepaper/QBI%204-25-05%20WSJ%20article%20with%20link.pdf [accessed 25 November 2010]

TechSmith Corporation, 2010. *MORAE usability testing and market research software.* [online] Available at: http://www.techsmith.com/morae.asp [Accessed on October, 2010]

Tidwell, J. 2006. *Designing Interfaces*. O'Reilly Media Inc.

Tullis, T. 2008. *Measuring the user experience collecting, analyzing and presenting usability metrics.* [online] Available at:

http://proxy.lib.chalmers.se/login?url=http://www.sciencedirect.com/science/book/9780123735584 [Acessed on November, 2010].

Urban-Lurain, M., 1996. *Intelligent Tutoring Systems: An Historic Review in the Context of the Development of Artificial Intelligence and Educational Psychology.* [online] Available at: http://www.cse.msu.edu/rgroups/cse101/ITS/its.htm [Accessed on 5 October 2010]

Vora, 2009. Web application design patterns. Amsterdam ; Boston : Morgan Kaufmann Publishers/Elsevier.

Wald, N., 2001. The definition of screening. *Journal of Medical Screening* 2001;**8**: pp.1-1

Wanhainen A., 2007. Massundersökning av kroppspulsådern räddar liv. [Internet] Available at:[http://www.akademiska.se/templates/page____33462.aspx]. [Accessed on 20 Mars 2010].

Whitton, N. 2010. *Learning with Digital Games - a practical guide to engaging students in higher education.* Routledge, New York.

Wiel-Hagberg, E., 2008. Livsavgörande screening erbjöds 65-åriga män. Borås Tidning, [internet] Available at: http://www.bt.se/nyheter/boras/livsavgorande-screening-erbjuds-65-ariga-man%28884359%29.gm [Accessed 25 April 2010].

World of Warcraft, 2010. [online] Available at: http://www.worldofwarcraft.com/index.xml [accessed on August 2010].

Wroblewski, L. 2005. *Web Application Form Design.* [online] Available at: http://www.lukew.com/resources/articles/web_forms.html [accessed on 20 October 2010].

Yaman, D., 2001. Why games work? [online] Available at: http://www.learningware.com/LearningCenter/WhitePaper1.html [accessed 26 November 2010]

List of figures

List of tables

Appendices

1. Persona
Erica, 45
Clinical Laboratory Scientist

Erica has been working within health care for 20 years and she feels confident with all her tasks at work.

Her primary job is to examine vessels by the ultrasound. During her long work practice she tested many ultrasound machines models and now she is looking forward to work on the newest machine with voice command.

Once every two weeks she works with aorta screening. This systematic examination is a five years long project, Erica feels excited to be part of it as she feels she is doing an important contribution. Her hospital joined the screening project two years ago and now, after many examinations, she feels confident about most of analysis and measurements she does.

Erica always works with a workmate, the company of another competent person in the examination room makes her sure that the work is done properly. *"It is better with two pairs of eyes to analyse difficult, troublesome images and find the vessel edge"*, she says.

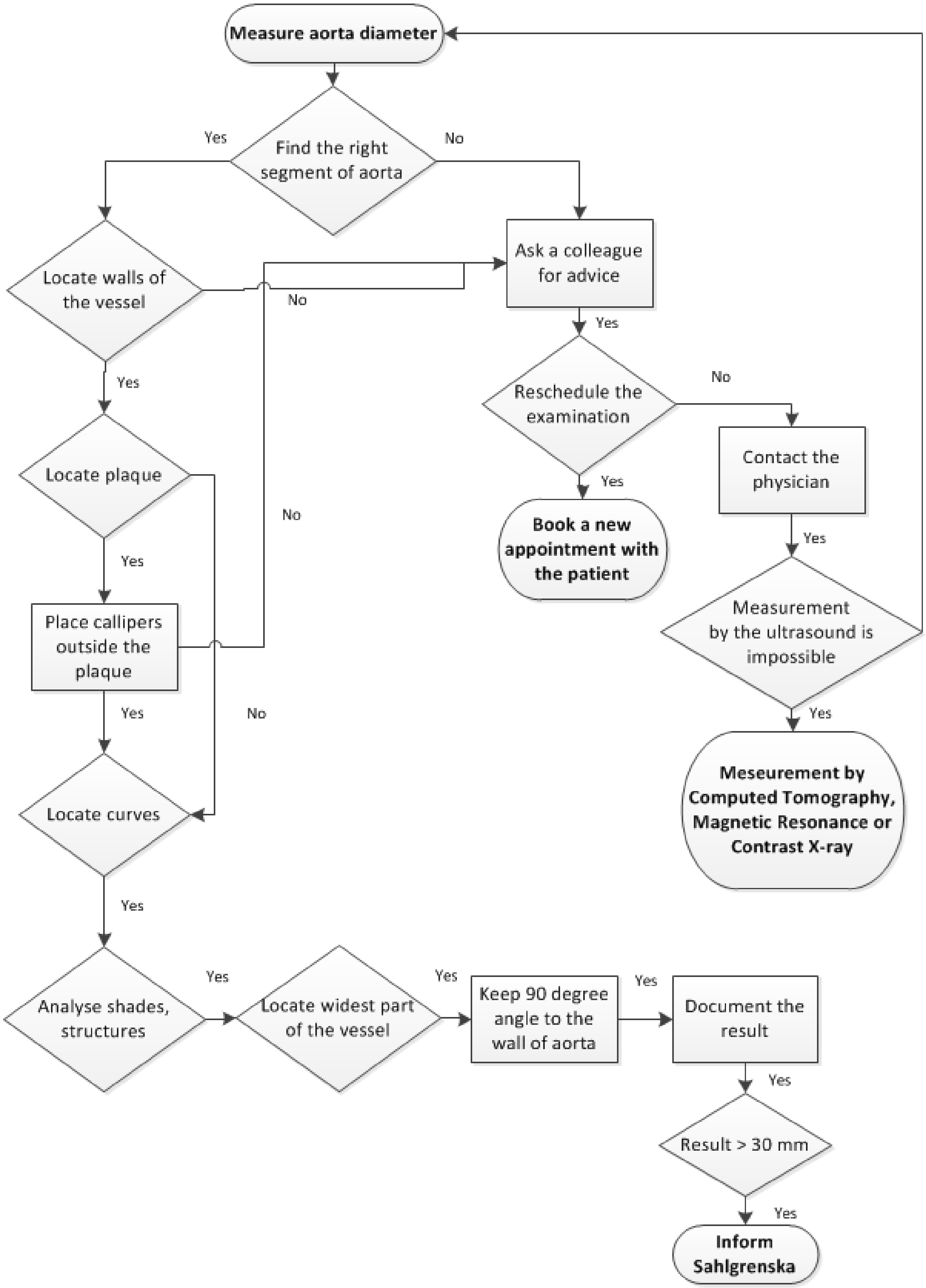
The worst case scenario would be to conduct measurements on the same person twice within a short period of time and get different results. That would be an evidence that one of measurements was wrong, this is a situation Erica and her colleagues want to avoid. She is aware that her measurement can save people's lives and she wants to measure as accurate as possible.

Erica feels confident about her measurement technique as the chief physician saw her working and he approved her work method. Sometimes however she feels that she would like to train by her own, too. According to her, watching and measuring challenging images would be valuable. Erica wishes she could train measuring tricky images with lime (kalken) and plaque on the walls of vessels as well as snaky, bent aortas, where the angle of measuring line needs to be assessed correctly. Erica wants to develop her skills. To learn more about e.g. vessels she accesses video sequences available on the Intranet and watches them and analyses carefully. She appreciates possibilities the computer programs give but she doesn't like to use mouse due to the wrist pain she has.

In the examination room there is an aorta screening instruction written by a chief physician, but Erica thinks it is a bit too simplistic and general. When the image is difficult, when vessel has jagged edges, shadows, darkness or other organs, gas in the way then the examination takes more time, sometimes Erica asks the patient to return another day or asks a colleague and the chief physician for advice. In rare cases they decide that a computed tomography examination has to be done instead of the ultrasound examination. Erica would like to train how to handle those “impossible” cases, she thinks it is important to know when to decide to give up the ultrasound examination and do something else instead. The screening work must go quickly, "*we only have 2-3 minutes for each patient and the decision must be taken rapidly"*, she says.

Erica remembers that in the beginning of screening work storing the result was challenging as it needs to be documented in both local and regional archive and marked with a special code. Writing the several-letter code and result is a bit tricky since it has to be very exact.

2. Workflow model


Measure aorta diameter
Find the right segment of aorta
Yes
No
Locate walls of the vessel
Ask a colleague for advice
No
Yes
Yes
Reschedule the examination
No
Locate plaque
Yes
Contact the physician
No
Yes
Book a new appointment with the patient
Yes
Place callipers outside the plaque
Measurement by the ultrasound is impossible
No
Yes
No
Yes
Meseurement by Computed Tomography, Magnetic Resonance or Contrast X-ray
Locate curves
Yes
Analyse shades, structures
Yes
Locate widest part of the vessel
Yes
Keep 90 degree angle to the wall of aorta
Yes
Document the result
Yes
Result > 30 mm
Yes
Inform Sahlgrenska

3. Persona based context scenario and requirements

1. Once every six months Erica accesses a special web site to train a measurement of aorta diameter. She is sitting in her workroom, in front of a computer.
She dedicates around 40 minutes to run the training application. The same amount of time it would take to conduct a real examination. The training of aorta diameter measurement is a part of quality optimization work that everyone needs to do regularly.
2. Erica reserved time for this learning activity in her schedule, her mobile phone however is on in case any urgent situation would occur. She runs the training application on different computers as she works in different sections of the hospital, depending on her schedule.
3. Erica starts the training program. She sees the result of her previous learning session and decides to do better this time. The first video sequence appears on the screen and Erica analyses it. The moving image provides much information about other structures lying around the aorta, as the sequence proceeds, she distinguishes walls of the veins from aorta's walls, she sees what white areas are shades and plaques. Erica recognises what movements are typical for gas or intestines.
4. When Erica feels she has enough information about aorta and adjacent structures she stops the sequence at the most suitable image and starts measuring. She places the callipers and can see the length in millimetres displayed. She gets an immediate feedback on the placement of callipers. Beside her callipers the correct callipers are shown in a distinctive colour. She sees clearly that she didn't place callipers in the right position. Program informs Erica what mistakes she did and how she can avoid them in the future.
5. She continues to view next sequence. Erica freezes the video and measures diameter, the application presents the result, this time she did a good job, her calliper are covered with a response callipers. The program confirms she did a correct measurement.
6. After accessing the last available video sequence Erica is informed about her result statistics, she did better this time and hopes to get even better after next session.

4. Persona-based key path scenario

Creating an account. Erica clicks the create new account button, she types her name and surname, then she chooses from the practice menu “experienced”, from occupational group menu “clinical laboratory scientist”, in the hospital menu she marks “Sahlgrenska”. The system creates a username out of entered data. Erica is asked to type a password she would like to use when accessing the application. She is asked to create a password of certain length and containing three digits.

Logging in. Erica logs in to the application. She is asked for a username and a password (användarnamn och lösenord). As she types wrong password, the error message is displayed: "Hoppsan! nu blev det lite fel. Det ifyllda användarnamnet och lösenordet är inte registrerat" "Observera du måste skilja på stora och små bokstäver, lösenordet ska bestå av mist sju tecken. Förskök igen". The message is written in red colour and proceded by a small, red error icon. Erica's second attempt to log in went just fine and she can now see the starting home page of the application. The first she can see is a diagram containing summary of her previous results, she can access the previous exercises or see details about errors, she may also look at the rules of the game, guidelines for measuring aortic diameter as well as a text explaining the application functionality and the interface.

Starting a new exercise. Erica wants to start a new exercise exercise as soon as possible. She chooses to push the prominent start-new-exercise button and the first sequence is displayed in the centre of the stage. On the left side of the screen that displays sequences Erica can see the Feedback panel with a short introduction text that says she needs to either push the ENTER button to start the sequence or click on the video player's screen. On the right side of the video player's screen she can see a small multiples presenting coming sequences. The uppermost square in the column represents a sequence that is currently displayed on the screen. Each small multiple provides information about what number in the exercise the sequence has as well as a still image from the sequence.

Choosing a frame from the video sequence. Erica views the sequence and decides to stop it at frame number 15, she believes that the chosen frame is the most appropriate for measurement. As she pushes the pause button on the video player panel, the feedback panel displays a text message informing Erica that she chose to measure on frame placed in the green area and that the correct frame is placed in the yellow area. Erica manipulates slider of the video player, goes to recommended yellow area and views that section of the movie. As she stops the video sequence again the feedback panel displays a text message informing that chosen frame was the right frame to measure on and that she can start placing electronic callipers within the image.

Placing callipers. Erica notices that a mouse cursor is now changed and that it has the shape of a cross hair cursor. She places the first calliper at the wall of the vessel and then the second, she looks at the image carefully and believes that the first calliper should be relocated a bit up. Erica picks the first calliper up by clicking it with the mouse cursor and places it at the desired location. She feels satisfied with the positioning of callipers and confirms her work by pressing the ENTER button on her keyboard. The video player's screen displays number of points earned. A summary of Erica's outcome is instantly presented to her in the feedback panel. Feedback panel consists now of three titled sections displaying type of conducted errors, measured diameter versus correct diameter (gold standard) and number points earned. After showing number of earned points the video player's screen displays the second sequence. As Erica feels she read all information from the feedback panel, she presses ENTER button on her keyboard and the feedback panel returns to its initial state, the outcome summary disappears and the second sequence is played in the video player's screen.

End of exercise. As Erica completes the last sequence of the exercise, the feedback on this sequence is presented to her in the feedback panel. In the area where the video player and small multiples were placed the summary of overall results from the exercise is displayed.

Summary of overall results. Erica sees her own achievements and how these were changing over time. The outcome of each exercise is represented in form of a square placed within a coordinate system. As Erica clicks on the square it becomes bigger and displayed in a distinctive colour, a text is displayed over the square, saying what exercise the square is representing, how many points Erica received in that exercise and which of four errors she committed in that exercise. The x-axis informs Erica what month and year she worked with each exercise, a horizontal slider lets Erica move forward in the timeline and view all exercises she has ever completed. The y-axis instead, displays percentage of the measurement she has done correctly, I.e the accuracy.

Scoring system. Every positive outcome of Erica is rewarded with scores. The system traces number times Erica attempts to stop the video sequence on the right frame for conducting measurement. If she found the correct frame to measure on at the first attempt she is rewarded with 2 points, if she was successful at the second attempt she receives 1 point, otherwise no points are given to the player. The system keeps track of the difference between Erica's obtained result and the gold standard. If the difference between these two is lower or even to 5% the player is rewarded with 10 points, if the difference is higher than 5% but lower or even to 7% the player receives 8 points, if the difference is higher than 7% but lower or even to 10% the player receives 5 points, otherwise, i.e. if the difference is higher than 10% no points are given to the player. The last value that determines player's overall result is number of errors committed. If Erica completed a sequence without any of four errors she is rewarded with 4 points, if only one error is committed the player receives 3 points, two errors give 2 points, a player with three errors gets only 1 point and four errors mean no points at all.

5. Cognitive walkthrough

Tasks

1. Create an account.
2. Learn about measurement principles.
3. Comment an image.
4. Start a new game.
5. Measure aorta diameter.
6. Control how much own measurement differs from Gold Standard
7. Control what errors the user has done.
8. Control how own measurement accuracy has developed over the time.
9. Control how experienced clinical laboratory scientists tend to measure, see how your colleagues measure.
10. Check the result of game 4 that was played previously.
11. End the game.

UA – user action
SD – system display

Task 1. Create an account.
UA 1: click ”skapa ett nytt konto” button in lower part of the page
SD 1: open the wizard

UA 2: fill out forms collecting personal information, experience and password
SD 2: display error messages or display the following page of the wizard. Once the password part is completed display a ”create an account button”.

UA 3: click the ”skapa konto” button
SD 3: display a rules of the game page

Task 2. Look into guidelines and principles concerning vessel measurement
UA 1: click ”Mätprinciples” link in the left menu
SD 1: display ”Mätprinciper” page

Task 3. Comment an image
UA 1: click the ”Bildanalys” button
SD 1: display the ”Bildanalys” page

UA 2: type a name, write a comment and click the ”Skicka min kommentar” button
SD 2: display user's comment below the latest comment

Task 4. Start a new game
UA 1: click ”Spela” in main menu
SD 1: display the two entry points

UA 2: choose ”starta ett nytt spel”
SD 2: display the first video sequence and instructions

Task 5. Measure the aorta diameter
UA 1: start the video sequence
SD 1: play the video sequence, activate the ”kontrollera om bilden lämpar sig för mätning” button

UA 2: click the pause button

SD 2: pause the video sequence

UA 3: click the ”kontrollera om bilden lämpar sig för mätning” button
SD 3: display a written directions in the ”Instruktion” section and a graphic symbol in the ”Status” section(either a cross or a check)

UA 4: manipulate the slider
SD 4: display corresponding images

UA 5: click the ”kontrollera om bilden lämpar sig för mätning” button
SD 5: display a written success confirmation in the ”Instruktion” section and a graphic symbol in the ”Status” section (a check confirms that a right image was found), mouse cursor acts as a calliper

UA 6: place callipers along walls of the vessels
SD 6: activate ”kontrollera min mätning” button

UA 7: click ”kontrollera min mätning” button
SD 7: display written feedback in the ”Instruktion” section, display measurement accuracy and types of errors in the ”status” section, display the Gold Standard callipers in the image, activate the ”nästa patient” button
UA 8: click ”nästa patient”
SD 8: display the second video sequence and instructions

Task 6. Control how much own measurement differs from Gold Standard
UA 1: view the in writing and graphic feedback in section ”Instruktion”, ”Status” and on the image itself
SD 1: -

Task 7. Control what errors the user has done.
UA 1: view the graphic feedback in section ”Status”
SD 1: -

Task 8. Control how own measurement accuracy has developed over the time.
UA 1: click the ”spara och avsluta spel” button
SD 1: display page with two entry points

UA 2: click ”Resultat” in main menu
SD 2: display user's diagram

Task 9. Control how experienced clinical laboratory scientists tend to measure, see how your colleagues measure.
UA 1: click ”Resultat” in main menu
SD 1: display user's diagram

UA 2: select BMA in the profession section
SD 2: display nothing

UA 3: select also all regions
SD 3: display nothing

UA 4: select even experienced

SD 4: display a curve showing an average measurement accuracy of all clinical laboratory scientists, in all regions who are defined as experienced

UA 5: select ”arbetskamraternas resultat”
SD 5: display additionally one curve showing a measurement accuracy of a colleague

Task 10. Check the result of game 4 that was played previously
UA 1: click the fourth circle in the curve
SD 1: display results of game 4 in the game summary section in the lower part of the page

Task 11. End of game
UA 1: click ”visa resultat” button
SD 1: display results of the current game in the game summary section in the lower part of the page

6. Results of heuristic evaluation

Results of heuristic evaluation – unique issues found by five evaluators.		
Design	Unique issues	
Design 1	Evaluator 1	
	POSITIVE FINDING: filtering the results in the results section was understood well.	
	Severity rating	Unique issues
	medium	1. When creating a new account it was pointed that list of options was too long, splitting the list into several screens was suggested.
	high	2. Evaluator could not come back to the starting site from site where a new account was created.
	low	3. The error message displayed when logging in with the wrong password was interpreted as impolite.
	high	4. Save exercise button made user doubt whether save button exits the exercise, too.
	high	5. The user examines 30 patients but the system displays 30 sequences. Using both terms to describe the same thing was seen as confusing.
	medium	6. The content of “Hem” was seen as unnecessary and repeating the content of “Spelet”.
	high	7. The content of “Spelet” was unclear, participant expected the game there not the rules and descriptions.
	high	8. Participant was confused when intended to start the game, navigated to the section “Spelet” instead of the section “Övningar”.
	low	9. The wording used in the instruction panel should be changed.
Design 2	Evaluator 2 and 3	
	Severity rating	Unique issues

Results of heuristic evaluation – unique issues found by five evaluators.		
	medium	1. Evaluator suggested that the way of displaying coming sequences is confusing and suggested that a progress bar would deliver the same information in understandable way.
	high	2. The way to display errors is confusing, participant was looking for a summary of errors committed.
	high	3. Evaluator didn't understand the why the points are given, suggested finding better way of providing summary of points given.
	low	4. Evaluator didn't understand what exercise is the current exercise.
	high	5. Evaluator could not come back to the starting site from site where a new account was created.
	low	6. Evaluator can't see any visual clue informing that it is possible to measure in the video sequence, wants a clear way to show that one image was chosen.
	low	7. Evaluator can't see how long time it took to measure the aorta diameter, suggested using a timer.
	medium	8. Evaluator doesn't understand the difference between an exercise and a patient.
Design 3	Evaluator 4 and 5	
	Severity rating	Unique issue
	low	1. The play button was too small.
	low	2. The text in the instruction should be completed with "press play alternatively Enter button."
	medium	3. Evaluator didn't fully understand the result summary.
	medium	4. The diagram presenting results should be completed with a panel showing details of the selected exercise.
	low	5. Presentation of golden standard result and user's result was seen as hard to read.
	low	6. Evaluator expected to be able to see own results in the "Hem" section, directly after entering the application.
	high	7. Evaluator was confused because terms game and exercise were used to denote the same thing.
	low	8. Evaluator couldn't see what exercise is currently done.

7. Tasks for heuristic evaluation and for paper prototype test

Funktion	Uppgift
Skapandet av ett konto	Skapa nytt konto
Inloggning in i webbapplikationen	Logga in

Funktion	Uppgift
Genomförandet av ett spel	Starta ett nytt spel
Feedback på om användaren väljer att mäta I rätt bild	Hitta en bild som är lämpligast att mäta i
Mätning av aorta diameter I ultraljud bilder	Mät diametern på aortan
Visa antalet poäng efter avslutad patient	Se hur många poäng du fick för mätning på patient 1
Visa vilka typer av fel användaren begick I en undersökning av en patient	Kontrollera vilka fel du gjorde när du undersökte patient 1
Presentera mätnoggrannhet efter avslutad mätning på en enskild patient	Kontrollera hur mycket ditt resultat av en undersökning av patient 1 skiljer sig från Gold Standard (hur noggrant du mätte)
Presentera mätnoggrannhet efter avslutat spel	Kontrollera hur mycket ditt resultat från spel 4 skiljer sig från Gold Standard (hur noggrant du mätte)
Visa antalet poäng efter avslutat spel	Kontrollera hur många poäng du vann i spel 4
Visa vilka typer av fel användaren begick i ett spel	Kontrollera vilka fel du begick i spel 4
Presentera spelets regler	Ta reda på spelets regler
Presentera alla deltagarnas resultat (visas som ett medelvärde – en kurva)	Kontrollera hur erfarna sjuksköterskor brukar ligga till
Presentera resultat av alla sjukhus	Kontrollera vilket sjukhus är bäst
Visa varje användarens resultatförändring över tiden	Se hur noggrannhet av dina mätningar har utvecklat under senaste året.
Presentera riktlinjer och principer kring mätning av diameter av aorta	Hitta information om hur du ska mäta aortadiameter
Publicering av ett resultat	Publicera ditt bästa resultat
Sparandet av ett oavslutat spel	Du behövs på en annan avdelning; spara dina resultat och avsluta spel du jobbar på
Återuppta ett oavslutat spel	Återuppta ett oavslutat spel

8. Usability test tasks

Uppgifter för test 1

1. Nytt användarkonto Skapa ett nytt konto.
2. När får man ett guldmärke? Genom att läsa informationen på hela webbsidan ta reda på hur noggrant man behöver mäta för att få ett guldmärke och titeln ”Mästare i aortascreening”.
3. Spela ett spel Nu känner du till spelets regler och kan börja spela. Spela ett spel och när du spelat klart berätta vilket märke du vann.
4. Vanligaste fel Se till att du befinner dig i Resultat och ta reda på vilket var ditt vanligaste fel som du gjorde i spel som du spelade nyss?
5. Hur noggrant mäter erfarna BMA:er? Ta reda på hur noggrant erfarna BMA:er mäter.

Uppgifter för test 2

1. Inloggning Logga ut och sedan logga in.
2. Vad är tredje mätprincipen? Ta reda på den tredje mätprincipen.
3. Spela ett spel Spela ett spel och när du spelat klart berätta vilket märke du vann.
4. Val av rätt bild för mätning Se till att du befinner dig i Resultat och ta reda på hur många försök behövde du för att hitta rätt bild för mätning?
5. Hur noggrant mäter erfarna sjuksköterskor? Ta reda på hur noggrant erfarna sjuksköterskor mäter.

9. Stakeholder interviews

Intervju 1

@ Recorded activity title: Interview
@ Recorded activity date: 20100505
@ Anonymized: Yes
@ Access: Research
@ Duration: 00:40:12
@ Participant: M
@ Recorder(s): Weronika Tancredi
@ Transcription name: Interview 1
@ Transcriber(s): Weronika Tancredi
@ Transcription date(s): 20100506
@ Transcribed segments: Almost All
@ Transcription system: GTS6
@ Time coding: Yes
@ Section: Arbetsuppgifter
@ Section: Utbildning, kompetensutveckling inom vården med hjälp av BFR
@ Section: Applikation
@ Section: Generellt
@ Comment: Only the relevant material was transcribed

§ START
§ Arbetsuppgifter #00:00:42
W: Hur länge har du arbetat inom Bild- och Funktionsregistret?
M: sen 2004, då var jag IT-ansvarig på Sahlgrenska, vi fick då frågor om DRA, som det här projektet heter, Distribuerat Radiologiskt Arbetsflöde. Sen hade vi lite synpunkter på det så tyckte min chef att man kan inte bara säga att någonting är dåligt så måste man bidra o(...) och så hamnade jag in i det.

W: Vilka arbetsuppgifter har du?
M: den titel jag har är systemförvaltare och det är allt ifrån att hantera det dagliga så att det snurrar till att man arrangerar inför framtiden och ta höjd för de önskemål om ändrat funktionalitet etc. Sen är det lite större, jag är ju också informationsansvarig för det genom att det har den kvalitén så som det ska vara och det har vi ett regelverk inom (...). Så det är inte bara att lagra ner men man måste då uppfylla det och uppfyller man inte det så tar vi inte emot det, då anser vi att då finns det inte materialet.
W: Är det så att ni kontrollerar enskilda sjukhus?
M: Vi kontrollerar alla. Så det är inte bara på sjukhusnivå utan det är den avdelningen som vill ansluta sig till BFR. Det de får först göra det är att ansluta sig till en testmiljö och skicka ner materialet på det de tänker producera och sen så går vi genom det mot regelverken och tittar på vad de uppfyller och inte uppfyller. Och så får de chansen att rätta till det, och så skickar de igen, och sen om de inte

kommer längre så går det till beslut för att göra undantag och då är det tidsatt att de måste åtgärda det innan den tiden, annars så kan de inte få lagra någonstans. Och det har en stor betydelse för i framtiden så tänker vi arbeta för att det som inte finns i BFR det kan man inte få betalt för. Eftersom vi sitter på ägarsida,(...) när vi köper vård så kan vi faktiskt säga vad är leveransen, dels är det att man ska behandla patienten dels för att man ska dokumentera och om dokumentationen inte finns i Bild- och Funktionsregistret så har man inte följt sin leverans och således får inte betalt. Så är det också att styra det för det kvalitativa perspektivet att man måste dokumentera på ett strukturerat och innehållsrikt sätt.
W: Vilken procent av vårdenheter som är anslutna till er?
M: Alla radiologiska verksamheter är anslutna, så där är det hundra procent och sen är det även de privata vårdgivare som bedriver radiologisk verksamhet, som regionen har avtal med, de har sitt krav att också lagra i BFR. Men sen har vi andra verksamheter som har tillkommit efter ett eget önskemål, sen har vi klinisk fysiologi, aortascreeningen, vi har kardiologin, mammografin, ja, det är massor och massor står i kö. Men eftersom det inte var tanken från början så har vi ingen organisation för detta så det är L. och jag som jobbar med detta, så det blir efter den bästa förmågan.
W: Vem använder ert material (era "kunder")? Från vilka sjukhus får ni materialet?
W: det har du svarat på
M: ja
§ Utbildning, kompetensutveckling inom vården med hjälp av BFR #00:05:10
W:Har BoF varit med och utvecklat eller bidragit till utveckling av material eller applikation för kompetensutveckling/utbildning av personal på sjukhus?

M: Ja, det har vi gjort, genom (...) att vi har grundmaterialet i BFR så kör vi ju varje år en kurs på IT-universitetet och där har vi (...) att ta fram applikationer som man kan återanvända det materialet vi har (...) det materialet vi får med det är egentligen det när du kommer som patient så undersöker jag dig och så ger jag dig ett svår på det och så arkiverar du det, det kommer så mycket som du arkiverar, men det är ju så att i hälsosjukvårdsprocessen så är ju det materialet vi tar inte bara för dig som individ och kan vi (an...) dig och sätta dig att du hade den frågeställning och andra populationen som hade den andra frågeställningen och så börjar man titta och så börjar man gå mot forskningshållet och se vad är orsaken till det, kan det vara miljöaspekter, kan det finnas gemensamma nämnare för den populationen etc, etc så att det har vi varit med om.
W: vem läser kursen?
M: det är studenter, vi ger de ett råmaterial och så har vi en inriktning och sen är det fria tankar, alla våra postrar som sitter där ute det är resultat från deras arbete. Då är det dels teoretiskt avsnitt i kursen som sedan leder till asp(...) period som sedan leder till en design och faktiskt framtagande av prototypen. (...)
W: och vad heter kursen
M:IT och samhället, tror jag den heter
W:Blir några av de förslag som de kommer fram till verklighet
M:o ja, vi har ett som ett företag som heter Synaps, och det är ett handdatorprojekt som också kom ifrån den kursen och sen ville de göra sitt exjobb och så anställde vi dem, och nu jobbar de och utvecklar applikationer. Egentligen är det att konsumera information i en handhållen dator, eller PDA. Så de finns verksamma och håller på (...) utanför regionen också.
W: Kan du ge ett exempel på ett konkret resultat av ert arbete.
M: Det är ju det med Synapsen är väldigt konkret och sen kommer det som du är inne på med aortascreening.

§ Applikation **#00:10:00**
W:Varför tror du att det behövs ett utbildningsprogram för operatörer av ultraljudsmaskiner som genomför aortascreening?

M: nu vet jag inte hur mycket ni har gått igenom om det här, man har ju stora kroppspulsåder den går ju ner (...) och så har det visat sig att främst män, kvinnor får det ungefär 10 år senare, främst män i ungefär 60, 65 års ålder har man hittat det åldersbandet får en sjukdom som gör att (...)ämnen slutar producera den mängd som det behövs. Så om man tänker sig en kroppspulsåder som en slang, och så är den lite elastisk, och så är det ju så att om du arbetar så får du ett högre tryck och då måste den vidga sig ut, (...) och när du får den här sjukdomen så blir det själva höljet (...) och det gör att då följer den av (...) följsamhet med tryckförändringar så i vissa fall får du förtunningar och då bildas det en bubbla och brister den så dör du med största sannolikhet. Och då har man sett det att män får det i ungefär 65 års ålder (...) kvinnor får det ungefär 10 år senare. Och eftersom med stort gott resultat kan operera folk i 65 års ålder och (...) så har man just nu börjat med männen bara (...) och då är det så att man mäter med ultraljud (...) och tittar på diametern på bukaorta. Det finns vissa toleranser man ser, en normal ska vara någonstans kring 30mm och sen uppåt är det olika span då tar man in de igen och sen är det direkt operation (...) och eftersom det är millimeter precision så någonstans kommer, (...) det beror hur du mäter eftersom ultraljud bygger på ultraljudsvågor som du skickar ner och som studsar upp så det är lite hur projektionen är hur mycket bukfett du har och det är mycket som kan avgöra. Sen så ser du den bilden och då är det markörer du sätter som maskinen mäter avstånd. Och när du kommer till gränsfallet då är det viktigt är du under eller är du över så det är därför behövs det kvalitetsoptimering, normering och det är därför jag ser att det här programmet skulle vara en hjälp för de här operatörerna att bli mer säkra på sin sak, de har ju inget svår om de mäter på dig så får de ju inget svår, jag ser ju vad jag gör men om jag har 10 kolleger till som mäter så kan vi få en standardavvikelse på hur vi talkar olika och asv. Så kan man göra en applikation där operatören kan sitta i lugn och ro och göra de här mätningarna på givna objekt då kan man se var man ligger och då kan man få en bättre förståelse om jag vinklar lite så (...) så att därför tror jag att det behövs.

W: Vilka grupper, tycker du, skulle ha nytta av en sådan utbildningsapplikation?
M: radiologer, BMA och sjuksköterskor som har intresse

W: Hur gick det till när ni bestämde er för att ni behövde detta program?
M: Ett dialogmöte med verksamhetsföreträdare där vi har träffat nästan alla som är delaktiga i den här verksamheten och lyssnade av de och hörde vad de var ute efter, och det är just det här att kunna hela tiden mäta sig själv i det. Dilemma med hela den aortascreeningen är just MÄTOSÄKERHETEN och det finns vetenskapliga studier i om du ska.. Det finns läkare i Uppsala som gjorde sin avhandling inom det här området, jag kommer inte ihåg.. Det var just det här om sjukdomsförloppet och hur man kunde diagnostisera det. Man har gjort mycket i England osv.

W: Finns det någon data som säger att det behovet finns?
M: Ja, om man gör en litteraturstudie så kommer det finnas. Där du ser att man diskuterar mycket hur det ska vara, hur man ska hålla proben, hur man ska mäta, så att det finns osäkerhet i det och många titta på göra den här golden standard. Detta (den framtida applikationen) kommer vara VGR:s golden standard, det är därför man vill ha det här programmet. Man har gjort lite mätningar och det är likadant. G. har gjort en presentation av deras (Uppsalas) arbete och då fanns de här avvikelserna, man ser utav grafer behovet att göra någonting. Sen kan jag inte säga någonting om att grafer kommer att se ut väldigt annorlunda efter att man har gjort detta, men det tror jag. För man blir mer komfortabel i hur man mäter, man ser direkt vad som påverkar eller inte. Så där finns det (...)

W: Vilka element bör, enligt dig, en utbildningsapplikation innehålla? (t.ex. feedback funktionalitet, likna ett spel, annat?)
M:Nån slags övningsintro och det behöver inte vara specifikt på hur applikationen sen kommer se ut utan det kan vara att man får en karta som man ska med markören mäta mellan två saker, två avstånd så att du får en känsla för det och kan sätta dig in att du vet från Göteborg till Växjö, om du bara har en karta över Sverige och inte städerna utsatta, så kan man kryssa i var Göteborg är, kryssa i var

Stockholm är, och så ser du det här vad du gjorde och var du ligger så många procent off från Göteborg, så många procent off från Stockholm och det här gör att hela mätresultatet blir så här mycket off. Bara för att man får lite känsla för det och att man kan koppla det till sin vardag och sen kan man gå in i det mer specifika då har man känslan med sig att det är inte så lätt även om alla känner till Sveriges geografi men man har en blindkarta bara att sätta städer det ger en lite uppfattning (...) Och det som du har skrivit här "likt ett spel" det är lite ytligt tycker jag, gärna lite med highscore eller förstår du "nu har du" och det behöver inte vara att man jämför sig med andra utan man får välja "vill du jämföra dig med andra?" och då kan du vara anonym och se var ligger jag till eller så kan du bara köra "du tävlar mot sig själv" på de objekt du har kört kan det vara individuellt, du kan även få ett samlat resultat, 1000 p säger vi, men då kan man säga att de 1000 p består av de här delmomenten under 5 undersökningar, här låg du väldigt bra, här gjorde du väldigt bra ifrån dig, för att annars så kan man bara säga 1000 – jättebra och då kan någonting (en del) faktiskt vara väldigt dåligt, då vet man inte vad man ska öva mer på sen, så att ngt sånt. Sen feedback review att den finns där att man kan få ngn som är väldigt duktig som har fått den rollen kan känna "det här känns riktigt bra" att man har en skicka-funktionen att någon som administrator som kan gå in på och göra en review av det, så att man kan få lite kritiska kommentarer på det: vad orsak-verkan är osv. Man ska även ha möjlighet till veckans case, de andra kan vara lite iterativa att man kör några veckor med samma mätmaterial, men sen så kan man lägga in ett case som är från materialet som blev producerat under föregående veckan, för då kommer många mäta på den och se hur stor avvikelse har vi, det blir lite real life case, då får vi också övergripande "var ligger vi" vad är golden standard från vecka till vecka, det är bara en tanke. Och sen att det baseras på spel, utmaningsdelen, gärna lite flashigt, så att man tänker: ju då jag vill komma upp till den platsen om man nu tittar och jämför sig med andra. Och kanske man ser med tiden hur har positionen förändrats, att man tänker: nu är jag nummer ett, men det kanske var bara denna omgången, du kanske har legat länge bra och dåligt sen, så att man kan se detta.

W: När och var, enligt dig, ska personalen använda den nya applikationen?
M: Så ofta de vill men minst en gång per termin, om man gör det med veckans case kommer folk använda det oftare, men jag tycker minst en gg per termin, det är den kvalitetsnormen det bör vara. Man får jättegärna göra det hemma men det som vi har krav på är att de ska göra det på arbetstid för det här ingår i kvalitetsoptimering av vår verksamhet och då är det en del av deras arbetsuppgifter. För det kommer sen det här att vi vet inte vad de har för hemmautrustning etc etc men vill de göra det ska de göra det men på jobbet ska det finnas möjlighet för det här.

W: Hur stor ska vara den skärm där applikationen visas?
M: Minst 24" skärm, gärna 30"
W: varför ska man ha så stora skärmar?
M (...) förstora för mätning kräver det
W: Skulle det funka på en PDA?
M: Man ska inte utesluta det men det är inte där man ska göra mätningen utan, jag ska inte svara så kategoriskt, man ska titta på det de arbetar på idag, det är där de ska köra applikationen på, så kör de på 17" idag så ska man köra tester (den nya applikationen) på 17" också och det ska vara så likt verkliga förhållande som möjligt, så ska jag svara. Sen vill man förfina så ska man göra det, den ska vara skalbar uppåt, men utgångsläget är så som de har det idag.

W: Använder anställda på sjukhuset webbaserade applikationer kopplade till PACS/patientjournaler där de kan utföra mätningar på ultraljudbilder? (dvs. utanför ultraljudsmaskinens gränssnitt)
M: Det finns det, det är inte så frekvent att man använder det för mätning på ultraljudsbilder därför att problemet är att när man sitter vid en ultraljudsmaskin så i att du har en geometri i kroppen (beroende på hur mycket bukfett patienten har får man olika vinklar, annorlunda projektion) så egentligen, om man ska vara riktigt noga så ska man lägga något referensobjekt så att vi vet hur stort exakt den är och kunna kalibrera mot det objektet så annars blir det ju projektionsavvikelserna som

kan (...)iffa i det du mäter. Därför är det väldigt svårt att gå från en applikation till en annan om du inte kan återkalibrera, dvs om jag tar en röntgenundersökning och så lägger en kula vid sidan, en stålkula som jag vet är 10mmi diameter och så vet jag att det är hälften av kropptjockleken, så ställer jag in det, så kan jag alltid kalibrera det med maskinen, det ska vara 10, så ökar sannolikheten om jag är inne i organen att det också är det objektet så det är lite svårt med webbaserat , om man inte har kalibreringsmöjligheten, men i BFR finns det ett webbgränssnitt som du kan gå in och mäta med. Men, det går an med röntgen och vissa applikationer där det inte är så stor efterfråga om noggrannhet, men det här är faktiskt på millimeter nivå så det får man ha i beaktandet att ha med den här kalibreringsfunktionen. Och det handlar om att normera sin mätning.
W: är största delen av mätningar gjort direkt på ultraljudsmaskinen
M: ja
W: och då räknar maskinen själv och kalibrerar
M: det finns lite olika metoder, vissa har det att du ska kalibrera mot något givet, och sen sätter du markören bara och då talar den om hur det blev
§ Generellt #00:32:25
W: Är det möjligt att automatisera mätning av aortadiameter på ultraljudbilder?
M: Nej, inte en chans i dagsläget, det måste ju komma en helt annan metodik. Därför att det är ett hantverk det här gör inte ven som helst utan det är de här specifika personerna. Andra tycker kanske annat men jag tror inte att det kommer under de kommande 5 års perioden. Finns inte en chans utan det är duktiga människor som kan göra det här.

W: Kan ett bättre gränssnitt (med fler, användbara funktioner och mer hänsyn till användaren) ge bättre mätning?
M: Ja, absolut, om man inte tar det bara som gränssnittet utan som en hel applikation. Övning ger färdighet, men visst du kan köra på väldigt många patienter men återigen om inte vi öppnar och tittar på bukaorta så, du få ingen feedback på det sättet. Men om vi kan ta patienten, det är där som du ska också prata med de du träffar, om man hittar någon som har väldigt stor diameter som de kommer att öppna och operera då är det ju bra att man kan på något sätt föra tillbaka det för då har man ett referensmaterial. Men det kan an göra på två olika sätt om man opererar, man kan antingen göra ett slutet ingrep, dvs man öppnar inte upp hela buken som man gör med traditionell kirurgi, utan man arbetar med endoskopisk kirurgi, man trycker in slangar i mindre ingångshålor och så opererar du med mekaniska verktyg inne, utifrån och inåt, men sen är det ju vissa som vi öppnar upp och har man mätt på en sådan så är det ju intressant att kunna mäta det under förutsättning att det medicinska ingreppet tillåter att man också kan ta den lilla tiden och mäta det fysiskt när man öppnar det och så kan man se. Och har man en sån patient där vi har det här gränssnittet så kanske det är 25 stycken som har mät på det, teoretiskt, och så få vi en normkurver över mätningen och sen när kirurgen öppnar och mäter det så kan vi se vad är avvikelsen då mot våran teoretiska norm och kontra den reella och då, då har vi tillfört något med stort värde som jag ser detta. Och sen är det så att repetition är kunskapens moder men då måste du få någonting tillbaka, annars så bara du gör och kan repetera ett fel om ingen återkopplar
W: gör man så här?
M: den här verksamheten startar vi i september och det är en pilot som vi ska köra i fem år för at utvärdera och se. Alla screeningsverksamheter ska ge samhälsekonomisk nytta. Vi kommer till en proaktiv sjukvård och hittar saker innan de ens brutit ut, planerat verksamhet är kvalitativt mycket bättre. Det finns risker med det vi gör men de är betydligt mindre, men det finns risker. Sen är det olika dilemmor: det här är inga patienter, de är medborgare som vi erbjuder en service, de kan säga nej till behandling. ( ) Vad har effekterna blivit, får man avlastat sjukvård. Vi måste optimera saker men att vi måste göra en switch till proaktiv sjukvård det är jag övertygat om. Proaktiviteten kan bryta den raka kostnadskurvan ()
W: vad heter projektet
M: AAS

§ END

Intervju 2
@ Recorded activity title: Interview
@ Recorded activity date: 20100512
@ Anonymized: Yes
@ Access: Research
@ Duration: 00:49:45
@ Participant: J
@ Recorder(s): Weronika Tancredi
@ Transcription name: Interview 2
@ Transcriber(s): Weronika Tancredi
@ Transcription date(s): 20100512
@ Transcribed segments: Almost All
@ Transcription system: GTS6
@ Time coding: Yes
@ Section: Arbetsuppgifter
@ Section: Utbildning, kompetensutveckling, personal
@ Section: Applikation
@ Section: frågeformulär
@ Comment: Only the relevant material was transcribed, when trancribing the words, the Standard Orthography was used, i.e. proper names, abbreviations/acronyms and punctuation were included.

§ START
§ Inledning #00:00:12
(...)
$J: (...) du behöver inte känna att du behöver bestämma metoden, utan att bestämma hur det ska se ut, men det kan vara bra om du ser om vi gör olika, om det är så
$W: men ska man göra en applikation då måste man följa riktlinjer, det måste vara något som är rätt
$J: jag har sett framför mig att vi skapar ett antal facit bilder, som inte innehåller någon mätning, såna bilder har vi redan, och sen att sätta sig och göra måttet i den bilden
$W: "göra måttet", det betyder att..
$J: Nu ska vi se om vi plockar fram en sån bild... så här ser det ut
$W: gör man mätning både i längdsnitt och i tvärsnitt
$J: vi mäter både och, och anledningen till att vi gör så här det är att om du kommer härifrån
@ < längdsnitt >
$J: så är det lättare att se om du ligger mitt i kärlen, om jag skär den här ute så får jag bara det avståndet, om man kommer från det hållet så är det lättare att se om man ligger i mitten. Om jag gör så, så är det svårare, och framför allt, om kärlet ligger så, så blir längden mycket längre, då blir det fel, därför är det en aning bättre att göra så.
$W: gör man alltid så här, med en längdsnitt?
$J: om de är stora så gör man alltid så att man tittar både och, för att om jag mäter så så verkar den mycket större, det är inte alltid de är helt runda. När de är sjuka så tittar man oftast på både håll, men man gör nästan alltid mätningen i längdsnitt där man har längden på aortan. Det ena trixet, det som vi inte kommer åt med det här programmet, det är att man är duktig på att ta bilden. Men när man fått bilden sen så där ser du att där är det ett plustecken och där är ett plustecken, där har man valt att mäta, därifrån och dit. Och det kan man träna sig fram till, de skillnaderna, där är bra att, även om ingen vet exakt vad är ett slutgiltigt facit, att man gör likadant, då blir det i alla fall samma fel hela tiden. Och det träningsprogrammet, man kanske har såna bilder, utan de här plustecknen och så får man föreslå själv då och så finns det en facitbild och kan den säga "lägg märken till att du bör lägga den där här" eller så finns det bara där och säger att den där är 42 mm (...) "du måste lagt den lite fel", någonting åt det hållet.

$W: Finns det något material, riktlinjer som säger hur man ska mäta ?
$J: Ja, riktlinjer har vi, Jag har studerat och försökt leta lite i litteraturen efter hur man ska mäta och då är det, en gammal variant, jag ritar här, så här ser en aorta ut, om man inte vet så mäter man det här avståndet
@ < se figur 1: den röda pilen >
$J: (...) man mäter från yttre delen av väggen (...) det här är liten skuggning av ekot men här börjar den, (...) det här heter (lidi..) alltså början på kärlväggen till början på (...)väggen. (...) Så här gör man på alla kärl, (...) men det är på ett rent och friskt kärl, så när man har ett sjukt kärl så blir det plötsligt lite svårare att veta vart man ska sätta kanterna någonstans, och då finns det de här riktlinjerna men det är lite svårare att bestämma sig för var själva kanten på vägen egentligen går och det är det som kräver lite träning att bestämma sig var man ska mäta någonstans på en sjuk aorta, var man ska mäta på en frisk är inte lika svårt. Då har vi gjort ett test här, vi satte en radiolog mot våra mest erfarna BMA som körde ett antal sjuka patienter här, för en månad sen, och såg vi att de skilde sig, vi har inte räknat färdigt än men de skiljer sig i ca 7% i snitt, det skiljer ett par millimeter, det kanske inte är hela världen de 2-3 mm, det kanske bara är deras teknik, det kanske är bildframställningen men det kanske är mättekninken och det har vi inte hunnit titta på.
$W: Den undersökning som ni gjorde, är det möjligt att titta på den?
$J: Ja, du kan få den sen, det är en BMA som håller på med sitt examensarbete, men det är klart att det ska ihop.
§ Arbetsuppgifter #00:12:05
$W: Då börjar jag med den första frågan: vilka arbetsuppgifter har du?
$J: Den här sektionen som gör kärlundersökningar; kärl- och njursektion, heter den, där är jag ansvarig för att de som jobbar på den kan det de gör, jag har inte personal ansvar, jag är inte chef för de men jag är den medicinske ledaren på den sektionen. Det är jag som ansvarar för att det finns bra riktlinjer och en del kan många saker bättre än jag kan så jag ser till att det blir rätt, att det fungerar på korrekt sätt på sektionen ur medicinsk synvinkel. Sen personal, löner, placering det finns det en annan chef för. Medicinska ledarskap, kanske.
$ Utbildning, kompetensutveckling, personal #00:13:41
$W: Hur vidareutbildas personalen på Sahlgrenska? På sjukhus i Sverige? I världen?
$J: Pratar vi allmänt eller aorta?
$W: Vi kan begränsa det till aorta.
$J: Jag tror att vi har ett ganska bra sätt att jobba på Sahlgrenska. Det kommer du märka om du går på de olika ställen att man kan göra aorta undersökningar på olika sätt, man kan jobba ensam eller man kan vara två och två och hos oss jobbar man två och två och det tycker jag har en stor fördel för då kan man sätta någon mer erfaren med en lite nyare och då lär de automatiskt upp varann, de kan också prata med varann om hur man ska hitta bilden, hur man ska mäta, i det vardagliga arbete (...) men de flesta ställen jobbar inte så, där kör man ensam. Och det är egentligen där, jag tror, man har största nytta av ett stödprogram, om man jobbar två och två så påverkas man av varann, byter man parkonstelationen så håller man ihop gruppen då alla försöker jobba likadant, man håller kunskapen i gruppen, inbillar jag mig. Men jobbar alla alltid ensamma så glider man ifrån varandra, utan att veta om det.
$W: Har ni alltid jobbat två och två?
$J: Vi har jobbat med det här bara i två år men redan från början bestämde vi två och två.
$W: Vilken utbildning fick personalen innan de började arbeta med aorta screening?
$J:Ja, vi satt tillsammans då, jag gjorde den här litteraturgenomgången för att se hur mycket det står i böcker och artiklar om hur man ska göra, men det står väldigt lite, sen var personalen på studiebesök i Uppsala som är landets ledande aortascreeningställe, de var först med detta. Och sen några av de som gör aortascreening är vana att undersöka bukkärl, benkärl så de har sett massor med oartor, även om de inte har sysslat med just rent screening då så är de vana vid att mäta kärlen. Vi började de fyra första och minst två av de var erfarna bukkärlsundersökare, sen hade vi några som var duktiga på ultraljud men inte hade kört så mycket bukkärl men de blev snabbt upplärda och sen tog vi två till, så att vi lärde upp internt.

$W: Hur många är de nu?
$J: Jag tror de är åtta nu.
$W: Och de arbetar heltid med det här?
$J: Nej, utan en förmiddag i veckan och då är det två av de som kör då. Så det är bara en gång i månaden de kör
$W: och hur många av de är BMI?
$J: allihopa
$W: Finns det en strategi eller ett program för kompetensutveckling för radiologer, sjuksköterskor, BMA? Vilket material används? Använder man "problembaserad inlärning" eller uppmuntrar man personalen att samla bilder i portfolio och reflektera över dem?
$J: Vi har satt oss ner och beskrivit hur vi ska göra och sen sprider det sig i det här gänget hur vi gör. Vi gjorde ett litet test förra året på en normal aorta och så gjorde vi det här testet mot radiologen, och skapat det här bildmaterial, syftet är att samla bilder och reflektera över dem, det är precis det vi tänkte göra i samband med det här examensarbete. När vi gjorde den testrundan med radiologen, de la ner bilder med och utan mätning på så att vi automatiskt skulle få ett antal bilder som går att använda till din reflektionsportfolio. Så att ett första råmaterial finns, sen får vi se. Den radiologen och de BMA som gjorde de bilderna ska träffas och titta på bilderna och avgöra vilka bilder som är bra och vilka som är dåliga och varför de är bra och dåliga och rensa lite och ge dig så bra bilder vi kan, det inget roligt att ge dåliga bilder att mäta på.
$W: Med dåliga bilder menar du att de är svåra att mäta i eller oskarpa?
$J: Ofta är det så att när det är svår aneurysm, vill du komma fram till den största (diametern), precis när du får den största så fryser du bilden och drar fram den, precis där du vill mäta är bilden bra, men här borta försvinner ju kärlet, det man helst skulle vilja ha är en fin bild av kärlet hela vägen men det blir det aldrig riktigt så. Den som står och tittar på bilden och tycker att den blir bra och så fryser man bilden och då har du tittat här hela tiden och det är där du ska mäta, så sen tittar du och släpper dina mätprickar och du bryr dig inte om här borta. Och sen kommer någon annan in och får den här bilden som inte varit med och undrar var har den här personen tänkt att jag ska mäta någonstans (...) Varje gång hjärtat slår så rör kärlet sig lite grann och regeln är att man ska mäta när det är allra störst, så i praktiken när du gör en undersökning så fryser du bilden och bläddrar en bit bakåt och lite framåt tills du hittar där rör sig kärlet och väljer en bild, då fryser du bara den bilden, när du sen plockar upp den så finns det inget att bläddra i utan det finns bara en bild. Då tänkte jag att vi kan kanske spara en längre sekvens som man kan bläddra i och om du har möjlighet när du gör ditt material att göra det möjligt att lägga in rörliga bilder lika väl som en fast bild så är det ju suveränt. Jag tror att det är ett bra sätt att göra det på.
$W: så att man själv får välja vilken bild som är bäst
$J: precis, då kan man låta den spela lite grann, en sekvens i två sekunder som går om och om igen och så kan du frysa och bläddra fram och tillbaka.
$W: Juste en till fråga angående riktlinjer, är de specifika för varje sjukhus, har Uppsala andra riktlinjer än Sahlgrenska?
$J: Jag tror inte det men det kan vara så, jag har inte sett alla andras riktlinjer, vi har inga gemen(...), jag tror att vi gör ungefär likadant om vi tittar på region. (...) Om det kommer ut det här utbildningsmaterialet, om det här blir bra, så kopplar vi det här till ett, då skapar vi också riktlinjer hur man ska göra, så jag ser det som ett parallell grej, vi säger från Sahlgrenska, eller i regionen, nu är det inte vi som bestämmer men om universitetssjukhuset hävdar att så här gör man och om man diskuterar det i regionen och om regionen är överens så ska det här utbildningsmaterialet vara ett sätt att öva sig på att göra så som det står i riktlinjerna.
$W: det finns inga centrala regler för det här
$J: nej, men de som är rankade högst, de som man frågar om man undrar det är Uppsala. När han kärnkirurgen från Uppsala var nere och vi berättade om det testgenomgången vi har haft då är det också han som föreslog att vi skulle ta fram det utbildningsmaterial kring detta (...) så gör vi det här bra så är det också någonting som Uppsala skulle vilja ha också då, och sen säljer vi det till England och Holland. Det svåra är ju att skillnaden mellan undersökaren kommer att bero mycket på det som

sitter i handen och ögat och den kommer vi inte att komma åt annat än med träning, sen finns det en mätvariation också och kan vi i alla fall ta bort den. Om vi har två fel så kan vi i alla fall reducera det ena felet, det är det det går ut på. Sen kommer alla inte alla få exakt samma mått på alla patienter men man kan minimera skillnaderna.
$W: och då pratar alla samma språk
$J: precis
$W: använder man "problembaserad inlärning"?
$J: Egentligen så är det exakt det vi gör, om vi sätter en ultraljudsvana BMA med en erfaren BMA som går på aortascreening som är varje patient ett problem och det löser de tillsammans så jag tycker att det här två mans arbete är väldigt bra.
$W: Du sa att man ska bygga ett gemensam portfolio men personalen individuellt samlar inga bilder?
$J: inte vad jag vet, vi har inget sånt system, alla aneurysm, alla som är minst 30mm och då lagras de i BFR, och då kan man plocka upp dem, kommer man ihåg ett personnummer så kan man det, de flesta av oss har intressanta fall som man följer. Men vi har inget systematiskt så annat än det jättesystematiska med att allt samlas i en egen låda i BFR (...)

$W: Finns det regler som säger hur mycket tid personalen ska ägna åt kvalitetsoptimering?
$J: Inte vad det gäller en enskild undersökning, det sägs att allt medicinsk personal på Sahlgrenska ska ägna 6 % av sin tid åt kompetensutveckling (...) annars så är det ju mitt jobb att se till att vi har ju mellan fem och tio olika sorters undersökningar på sektionen, man måste se till vilka undersökningar är i störst behov av utveckling just nu. Och just nu lägger vi tid på aortascreeningen med dig.
$W: Hur jobbar man med kvalitetsoptimeringen?
$J: Det är lite olika (...) dels försöker vi jobba likadant, att vi jobbar tillsammans på samma patient, i andra sammanhang jämför vi våra resultat med röntgen resultat , stämde det, dels jobbar vi mot kärnkirurger, det är rätt mycket, så egentligen är kvalitetssäkring någonting annat, man går på en rond, det är också kvalitetssäkring. Men det är en del av vårt jobb. Det måste vi göra.

$W: Hur ser utbildningen av den personal som mäter aorta diameter på ultraljudbilderna ut? Lär man sig av andra? Böcker? Videomaterial? Det har du svarat på.
$J: ja
$W:Hur många personer arbetar med aortascreening på Sahlgrenska sjukhuset? andra sjukhus?
J. jag tror det är åtta stycken
$W: och de andra ställen
$J: det vet jag inte
$W: ja
$W:Finns det något material tillgängligt om kompetensutveckling och utbildning av personal på sjukhus?
$J: riktlinjer finns
$W: Forskning kring metorder inom diagnostiken av pulsåderbråck - vad ska jag läsa? I Sverige, utomlands?
$J: Det är det att man jämför ultraljudsresultat med datortomografi och de metoder ger ofta en åning större diameter än ultraljud gör, men det tror jag att det ligger i att metoden är som den är, de mäter ofta lite längre utanför än vad vi gör och (...) det är det resonemang vad är det vi vill uppnå, vad är det det ska åstadkomma. vi vill hitta de som växer och då har man sagt att det finns en gräns på 30mm och den nu är satt med ultraljud då är det ultraljudsmätning som gäller. Och sen när man opererar så opererar man ofta på en CT?(P) utlåtande och i de nivåerna då är det CT?(P) måttet som ska gälla då och det är det något som man ska diskutera med våra lokala kärnkirurger för det är de som bestämmer när man ska operera och i vilken nivå som man ska operera. Det finns ju riktlinjer men de är olika lokalt vad man gör i Stockholm eller i Uppsala. Det kan du kolla vad Anders Wanhainen skriver. Annars så finns det en del studier kring varför man skall göra screening, (..) att man sparar liv och pengar vilket är jättebra. Men sen om man ska läsa hur man har kommit fram till, sen försöker jag läsa hur man gjorde sina ultraljud så är det knappt en mening, metoden är beskriven väldigt

sparsamt så man får leta bakåt, bakåt. Det är för att det inte är jättekomplicerat men det inte alltid står tydligt skrivet exakt hur man ska göra.
§ Applikation #00:35:42
$W:Varför tror du att det behövs ett utbildningsprogram för operatörer av ultraljudsmaskiner som genomför aortascreening?
$J: reducera mätfelet, så att alla mäter på samma sätt

$W: Finns det någon data som säger att det behovet finns?
$J: Nej, det är bara en känsla. Nej, om vi gör det här programmet så måste vi också visa att det behövs och våran tanke här i höst är att göra en sittning med ett antal aneurysm patienter så att fyra olika grupper mäter och sen några gör det här programmet och några gör det inte och så mäter man en gång till och då ser man hur mycket bättre blev de som gjorde det här programmet och det här måste ju gå, annars så behövs det ju inte. Det kanske räcker att man går tillsammans med varandra och får riktlinjer så att vi vet ju faktiskt inte att det behövs. Det är faktiskt så. Men om det behövs så är det ju faktiskt något nytt som kommer.

$W: Vilka grupper skulle ha nytta av en sådan utbildningsapplikation?
$J: De som skulle ha mest nytta det är de som är nya, nyast men de som är på väg att lära sig att göra detta, de som är i introduktionsskede, och så tror jag även de erfarna som en säkring av att alla mäter på samma sätt. Även om man har två, tre jätteerfarna men om de mäter olika så blir det ju olika, de kanske har lärt sig olika från början så blir det ett sätt att synkronisera mätsättet i regionen, målsättningen

$W: Hur gick det till när ni bestämde er för att ni behövde ett sådant program?
$J: Vi hade en träff om aortascreeningen i regionen och Anders Wanhainen från Uppsala var här nere och så beskrev vi vårat sätt att göra de här testerna och då pratade vi om att man kanske kunde ha nytta av att göra det här utbildningsprogram, och då sa vi att vi testar. Så det var en gemensam idé mellan oss och Uppsala.

$W: Vilka element bör, enligt dig, en träningsapplikation för mätning innehålla? (t.ex. ge användaren feedback, utmana användaren, annat?)
$J: Bilder där du själv ska kunna lägga in mätpunkterna och sen en feedback, antingen när du har lagt in punkt så får du..., idealet skulle vara att du lägger in punkt, du har en stor skärm så lägger du en punkt (...) sen lägger du in den hit och så kommer en facitbild här (...) och kan du titta aha jag fick åttiotre och det borde vara åttiofem. Och eventuellt om det finns en teaching point om det finns någon kommentar, en facit, t.ex. Notera att det vita uppe till höger inte är väggen utan det är en artefa(...) det kan alltså finnas som en liten kommentar man lägger just vid den bilden. Och särskild med rörliga bilder ska man kunna se om man har lagt de samma bilden om den här ligger skilda åt. En bildsekvens innehåller ett antal frames, ett antal bilder ovanpå varann så att man har två sekunder innehåller sextio bilder och så har du mätt p bild tjugoåtta men facitbilden är bild trettiotre, det är också en information att facit, de erfarna valde att mäta på en bild som ligger en halvsekund senare och då kan man jämföra det lite grann. Frågan är om man ska möjligheten att gå till nästa fall eller att gå tillbaka om mäta eller någonting sånt här.

$W: När och var, enligt dig, ska personalen använda den nya applikationen?
$J: Jag vet inte. Det beror på. Dels så ska vi försöka systematiskt utvärdera programmet från början och till det så kommer vi använda de flesta av vår personal men sen så kommer man använda det då och då, en gång om året eller vad det kan vara. Så kör man igenom trettio bilder eller att man gör det gemensamt i grupp kanske, en gång om året. De nya gör det när de är på väg in i verksamheten och alla andra kör det som en repetitionsgrej då och då, typ en gång om året.

$W: Hur stor ska vara den skärm där applikationen visas?

$J: det ska funka på en vanlig upplösning, eller att man ska ha en dubbel skärm så att du jobbar här och så får du en facit där, de flesta av oss har ju dubbla skärmar nu för tiden Allting ska digitaliseras. Man jobbar ju ofta med bilder här och journal där så att dubbla skärmar är egentligen väldigt vanligt, så det kan vara bra.

$W: Använder anställda på sjukhuset webbaserade applikationer kopplade till PACS/patientjournaler där de kan utföra mätningar på ultraljudbilder? (dvs. utanför ultraljudsmaskinens gränssnitt).
$J: Vi har arbetsstationer, de flesta mätningar görs i ultraljudsmaskinen, online då, kallar vi det för, medan man är uppkopplat på patienten, det görs det och sen läggs de in, i praktiken så, när är man nöjd, så sparar man bilden och sen sker det en automatisk överföring via kvalitetskontrollen till BFR och så tittar man aldrig mer på dem egentligen. Jag tittar ju aldrig på de trettio bilderna som gjordes idag så att de gjorde rätt, Så funkar det inte, utan det om det kommer frågor efteråt. men en liten egen journal bildföring så att man ska alltid kunna hitta det. Sen har vi offline stationer där vi kan på konkret arbete och plocka upp saker och mäta i de har vi inte använt jättemycket och just med det här aortascreeningen har vi svårt att komma åt det,vi har inte fått tillgång till aortalådan ännu, så det är precis nyss ordnat, det är massa regler kring det, eftersom alla i regionen lägger ner sina bilder i samma låda så måste alla i regionen godkänna att alla får titta på dem, så det är massa såna på skrifter som ska runt så att det ska bli formellt korrekt. Annars så använder vi det lite grann, de mesta görs i undersökningsrummet och sen görs det väldigt lite efterarbete som det är nu.
$W: Om man får upp en bild från BFR kan man mäta i bilden en gång till?
$J: Ja, det kan du göra men de saker som du satte tidigare sitter kvar. Jag har aldrig lyckats att ta bort den, finns det en mätning i bilden så är den fryst så att den sitter ihop, man kan inte ta bort det. Så det är därför vi gjort så här att vi sparat en tom bild, att det ska finnas en tom bild.

$W: Ligger problemet med mätning av aorta diameter huvudsakligen i att det är oklart hur man ska mäta (det saknas en metod/ en standard) eller i att gränssnitt på ultraljudsmaskinen är ett förvirrande och krångligt.
$J: jag tror att, det här är bara en gissning, det kommer vi att veta mer om när vi gjort det här programmet. För är det så att det här programmet inte gör någon skillnad så ligger ju hela felet hos den som kör, alltså i bildframställningen. Det ena är bildframställningen och det andra är mätningen på den framställda bilden och jag tror egentligen att det här är känsligare, här är det större variation, men jag tror att det finns tillräckligt stor variation i mätningen så att det ska vara värt att säkra den, om inte annat så ska det finnas tydliga riktlinjer så alla gör likadant och då har vi en stor styrka att vi vet om alla gör likadant och även om det visar sig sen att vi mäter 2mm fel så är det bara att översätta det till och förändra två mm. Men det vet vi egentligen inte, inte i någon större skala.
(...)

§ FRÅGEFORMULÄR #00:49:45

§ END #00:49:45

Intervju 3
@ Recorded activity title: Interview
@ Recorded activity date: 20100519
@ Anonymized: Yes
@ Access: Research
@ Duration: 00:28:00
@ Participant: G
@ Recorder(s): Weronika Tancredi
@ Transcription name: Interview 3
@ Transcriber(s): Weronika Tancredi
@ Transcription date(s): 20100519
@ Transcribed segments: Almost All
@ Transcription system: GTS6
@ Time coding: Yes
@ Section: Arbetsuppgifter
@ Section: Utbildning, kompetensutveckling, personal
@ Section: Applikation
@ Comment: Only the relevant material was transcribed, when trancribing the words, the Standard Orthography was used, i.e. proper names, abbreviations/acronyms and punctuation were included.

§ START
§ Arbetsuppgifter #00:00:10
$W: Vilka är dina arbetsuppgifter, vilken funktion har du i aortascreeningprogrammet?
$G: Framförallt forskning, delvis utbildning, planering, inte praktiskt i screening
$W: ingår framtagning av riktlinjer och att se till att personalen..
$G: kan det de ska göra
$W: Hur länge har du arbetat med screening, har du varit från början?
$G: Ja, jag har varit med från början
§Utbildning, kompetensutveckling, personal #00:01:11
$W: Hur vidareutbildas personalen på Sahlgrenska? På sjukhus i Sverige? I världen?
$G: Det sker egentligen ingen speciell utbildning annat än att man går bredvid
$W: Två och två?
$G: Två och två och av och till gemensamma gruppdiskussioner om kriterier och var man ska mäta och hur man ska mäta, men någon formell utbildning finns inte.
$W: Har man planerade möte där man diskuterar vad som har hänt och hur man mäter?
$G: Inte, vi har vanliga APT men inget dedicerat aortascreeningmöte, varje år försöker vi ordna Västra Götalandsregionen möte, i September, det handlar om allting och mätning är bara en liten del av det.
$W: Finns det en strategi eller ett program för kompetensutveckling för radiologer, sjuksköterskor, BMA? Vilket material används? Använder man "problembaserad inlärning" eller uppmuntrar man personalen att samla bilder i en portfolio och reflektera över dem?
$G: Tyvärr finns det inte det, utan det är det som du håller på, som är strategi, just nu håller vi på och inventerar problemet, det finns en exjobbare som gör mätningar på massa olika aorta diametrar för att illustrera behovet av träning, hon är snart klar nu i veckan.

$W: Finns det regler som säger hur mycket tid personalen ska ägna åt kvalitetsoptimering?
$G: Nej, det finns någon som sagt att i utveckling och utbildning så ska det vara 6 % av budgeten men den siffran vet jag inte var den kommer ifrån, om den är sann, det vågar jag inte svåra på än.
$W: Denna fråga hänger ihop med hur mycket tid tycker du de ska lägga på att träna
$G: De behöver inte träna så mycket, tror jag, de är jätteduktiga, det som eventuellt vi gör fel är att vi gör lite olika mellan person till person vi skulle behöva synka oss, samordna oss och kanske ha problemdiskussioner när det dyker upp problemfall
$W: så programmet skulle vara en plattform för att ta upp såna speciella fall?

$G: Ja, det tror jag och sen, när det kommer nytt folk så mäta på samma sätt så det är viktigt med material och formell träning, så att kommer det folk hit så lär vi de upp. Men de som är i programmet och redan gör det, de tror jag är duktiga. Men just för tillfället så tror jag att det finns en viss synkroniseringsmöjlighet i att vi mäter lite olika och uppfattar olika bilder på olika sätt.
$W: Jag förstår att det finns två delar som påverkar mätning, dels hur man mäter på de bilder som redan finns och dels hur man tar fram bilder så jag undrar om programmet ska ge möjlighet till träning av alla delar.
$G: Programmet ska ta hand om mätning, inte handhavandet men lite av det andra också. Vi kommer att ha en loop, vi lägger ner en filmsekvens som är kanske femtio bilder och då ska man välja att mäta på rätt bild. Vi kommer svepa genom kärlet, det vore bättre, inte en stillbild. Man ska visa en filmsnutt och stanna och så ska man välja vilken frame man vill använda till mätning och då kan en del av uppgiften bestå av att välja en rätt frame.
$W: Finns det videosekvenser sparade i Bild- och Funktionsregistret?
$G: Tyvärr, vi la inte ner dem, men de kommer i september att läggas ner.
$W: Samlar man bilder i portfolio?
$G: Inget organiserat, men jag har ett eget system ibland där jag kan markera upp, inte via BFR, jag kan inte markera de via BFR men jag kan markera de på andra ställen, intressanta case. Det är ett litet forskningssystem, då kan jag få upp dem om de har en kryss i en ruta och sen kan jag hitta det krysset och (...) det är ett sätt att markera att det här är en spännande patient, det finns inget annat sätt att göra det här. Folk ofta en kopia på en remiss och så sparar den, det finns en låda och så tittar de senare på vad som händer med patienten.
$W: Det skulle vara intressant att titta på det närmare.
$G: Ett sätt att markera en bild som intressant för undervisning vore bra.
$W: Hur ser utbildningen av den personal som mäter aorta diameter i ultraljudbilder ut? Lär man sig av andra, tittar man på en demonstration? Böcker? Videomaterial?
$G: Nej, ingenting alls.
$W: Man ska ha en facitbild och vem bestämmer vilken bild som är mätt på ett korrekt sätt?
$G: Det har vi gjort nu, vi resonerade så här att den som ska operera patienten är den som bestämmer vem som ska vara facit, vad är facitmätning. Kirurgerna med Urban Vindgren bestämt att vi ska jämföra vår ultraljudsmätning med en annan ultraljudsmätning, traditionellt så är det röntgenläkare som gör de här undersökningarna (aortamätningarna). Innan vi började med screening så var det enstaka mätningar som röntgenläkare gjorde, de har alltså längst erfarenhet av de här mätningarna och då definierade vi, tillsammans med kirurgerna, att nu ska vi ta en erfaren radiolog som golden standard och det har vi gjort.
$W: Hur vet radiologer hur de ska mäta, utgår de bara från den egna erfarenheten?
$G: Man läser böcker, försöker se vad andra gör, erfarenheter, men här på Sahlgrenska finns det ingen kvalitetssäkring av vad de gör. Och nu har vi jämfört oss med radiologerna, och det är det exjobbaren har gjort. Och då visade det sig att våra mätningar, för det första så skiljer de sig måttmässigt inte signifikant från radiologen. Om radiologen har 100%, the golden standard, så 95% av våra mätningar hamnar i plus minus 10% av vad radiologen gör.
$W: Är det bra?
$G: Det är någorlunda bra, acceptabelt. Då är det vissa som skiljer ut sig mycket, vissa har stor skillnad, andra liten skillnad. Det var några problemfall med i undersökningen, stora, konstiga kärl så blev det större spridning.
$W: Riktlinjerna, kan man titta på de?
$G: Metodbeskrivning tror jag att det finns, visst kan du det.
$W: Det var åtta anställda som jobbar med aortascreening på Sahlgrenska och de är BMA allihopa, stämmer det?
$G: Ja
$W: Forskning kring metoder inom diagnostiken av pulsåderbråck. I Sverige, utomlands?
$G: Det blir någon form av forskning det du gör just nu. Diagnostik är inte så jätte mycket forskning om, det gäller att ha hög kvalité. Vi var väldigt noga med att mäta det vi gör men det är bara en

kvalitetsgranskning internt. Det är inget jag skulle kunna kalla forskning. Om vi kan hitta ett sätt som gör att vi förbättrar mätmetoder i Västra Götaland systematiskt i aortascreeningprogram då skulle man kunna kalla det forskning och skulle kunna publicera det. Om vi hittar en utbildningsplattform som är webbaserad, till exempel, och som vi kan påvisa att den förbättrar diagnostiken, gör diagnostiken säkrare så finns det publikationsvärde. Sen forskar vi allmänt om vad händer med de här patienterna, de som har aneurysm, hur går det för dem, vilka sjukdomar har de och vad händer med dem. Uppföljning.

§ Applikation # 00:14:04

$W: Varför tror du att det behövs ett utbildningsprogram för operatörer av ultraljudsmaskiner som genomför aortascreening?

$G: Nu håller vi på ställa fram trettio bilder och vårt mål är att låta alla i regionen mäta på dessa bilder, samma bilder, jag tror att vi kommer få olika resultat på olika ställen, det betyder att folk mäter på olika sätt. Nu håller vi att framställa bilderna, men vi har problem med BFR, folk kan inte titta i bilderna så det är det som stoppar oss. (...) Bildmaterialet finns, det finns mätare i regionen som är intresserade av att mäta men vi kan inte visa bilderna. Applikationen behövs för att få folk att göra på samma sätt och mäta på samma strukturer, det är bra att få folk att göra lika fel överallt, jag tror inte vi kan hitta sanningen, för den är svår att hitta men alla skulle kunna göra lika fel, det är ett stort framsteg.

$W: Vilka grupper skulle ha nytta av en sådan utbildningsapplikation?

$G: läkare och BMA, de som kommer in nytt och även de som gör det och hållit på länge kan behöva fräscha upp sig och en gång om året checka av sig så att de inte glider iväg börjar mäta annorlunda. En kontinuerlig kontroll av kvalitén och specialträna nya personer. Aortascreening är inte genomfört i alla landsting ännu så de som nu ska starta upp det här står inför uppstart och då är det bra att de har ett system som de kan träna upp sig. Det skulle vara fantastiskt om de kunde logga in och enkelt ta reda på hur det här fungerar och testa sig mot sanningen.

$W: Hur gick det till när ni bestämde er för att ni behövde ett sådant program?

$G: Jag har länge tänkt att vi behövde bildbibliotek för att träna oss på, och nu när aortascreening dök upp så var det naturligt att träna på aortascreeningen. Det finns flera applikationer som man kan tänka sig för andra undersökningar, samma typ fast då blir det mer komplicerat. Vi gör undersökningar på njurar, halskärl och det skulle vara nyttigt att lägga in de i ett sånt system, men då blir det lite mer komplicerat än bara ett mått. Så det här är ett bra testsystem eftersom det är relativt enkelt.

$W: Finns det någon data som säger att programmet behövs?

$G: Nej, det finns ingen data alls, vi vet inte alls om vi är bra eller dåliga men vi försöker ta fram denna data, man kan se på det här exjobbet att otränade personer mäter mycket sämre än tränade personer (...). Vi hoppas att kunna visa att otränade personer som exponeras för programmet så ska de bli bättre. Det hoppas vi kunna visa i höst.

$W: Vilka element bör, enligt dig, en träningsapplikation för mätning innehålla? (t.ex. ge användaren feedback, utmana användaren, annat?)

$G: Dels så att det är naturlig miljö, att man känner att man jobbar i någorlunda samma system som man gör i vardagen, att man känner igen lite mätverktyg och ungefär samma saker. Sen är det väldigt bra om man på något sätt jämförs ens resultat, om inte online så efter avslutat test med någon form av referenspopulation, att man får feedback alltså, man får reda på var det här bra eller dåligt, vilka patienter var dåligt och feedback. Programmet ska ge feedback baserat på golden standard. Det ska vara helt självgående, däremot så en gång om åter så reviderar vi lägger till nya fall, nya kommentarer. Varje år eller varje halvår så skapar vi fem nya fall och kastar ut fem gamla. Det här systemet ska hållas uppdaterat så att man inte har massa gamla bilder från gamla, dåliga ultraljudsmaskiner, utan kontinuerligt ska det finnas en uppdatering. När individen går in så individen börja och sluta framför dator och inte behöva en feedback från en fysisk person om det ska bli framgångsrikt.

$W: Jag tänkte lite på samarbete, att man ska titta på varandras resultat, om man vill diskutera ett fall eller fråga varann

$G: man kan tänka sig att man sparar resultat anonymiserat, så som, till exempel yrkeskategori, tränad, otränad, undersökare nummer ett. När du börjar din session så får du ange vem du är, är du läkare, är du BMA, är du student, och sen får du säga har du jobbat jättemycket med detta eller ingenting alls eller mittemellan. Alla resultat kan sedan redovisas för de här grupperna, så får man t.ex. En student brukar klara sig så här bra, väldigt tränade personer brukar klara sig så här bra- möjlighet som en tänkt feedback.
$W: Kan det finnas element av spel, som t.ex. Utmaning, tävlingsmoment?
$G: Det kommer det bli om man redovisar resultaten. Tävlingsmoment uppstår om man säger sjukhus ett, sjukhus två, sjukhus tre. Du själv vet vem du är – jag vet att jag är sjukhus tre, men ingen annan vet det. Jag vet att jag är undersökare femtioåtta, men ingen annan vet det. Allting ska vara anonymt som går in (inte ens administratören ska veta det). Någonstans ska det finnas en kod vem gör detta så att man kan se att ingen fuskar, att en väldigt duktig person går in på studentsida och messar upp med data. I princip så tror jag att om det ska ha en stor användning så ska det vara väldigt anonymt. Man ska dock kunna tänka sig att kunna identifiera sig om de som går in i det vilja göra det själva, om de vill få en personlig feedback. Men annars så fungerar det bäst om individen själv får en feedback. Sen kan man använda det som en check, om du ska få göra den här undersökningen så måste ditt värde (genomsnittlig avvikelse) vara under 10%, då får du inte jobba med detta. Då blir det faktiskt körkort grej. Till exempel den medicinsk ansvarige kräver att man ska genomgå det här testet och vill också se resultatet och blir det godkänd så får man göra den här undersökningen annars så får man utbilda sig bättre tills man klarar testet.
$W: När och var, enligt dig, ska personalen använda den nya applikationen?
$G: Dels under upplärning, när man lär sig det här överhuvudtaget och sen kanske en gång om året, som kvalitetskontroll.
$W: Hur många bilder ska man mäta?
$G: ju fler desto bättre, för studenter så finns det ett enormt behov av såna grejer. Vi har en bildbank på hundra patienter som vi uppdaterar och håller fräsh med kommentarer, det är det vi har tänkt oss att lägga in hundra.
$W: När skulle de träna, ingår det i deras arbete
$G: Ja, på arbetstid.

$W: Hur stor ska vara den skärm där applikationen visas?
$G: Vanliga, SU datorer, sjukhusdatorer, inga fancy. Det blir mest genomslag om man kan köra det som en webblösning men om man kräver en arbetsstation, om det är tyngre grejer så kan man ordna det också men då blir det mindre genomslag. Jag vill att någon från Växjö kontaktar mig och säger kan inte vi mäta. Då kan jag hänvisa till denna hemsida och ge inloggningsuppgifter så får jag reda på var de ligger just nu. Det får inte vara helt publikt, men man kan tänka sig en liten reklamfilm med några exempelbilder som visar hur det funkar, men sen måste man logga in för att komma åt det fulla systemet.

$W: Använder anställda på sjukhuset webbaserade applikationer kopplade till PACS/patientjournaler där de kan utföra mätningar på ultraljudbilder? (dvs. utanför ultraljudmaskinens gränssnitt).
$G: Vi har ett egenutvecklat bildhanteringsprogram som heter *Web Adapt* för radiologer, där kan man mäta.
§ END #00:28:57