NTNU
Norwegian University of
Science and Technology

A Collaborative Tool for Medical Research

Heidi Johansen Heggelund

Master of Science in Informatics
Submission date: May 2011
Supervisor: Pieter Jelle Toussaint, IDI
Co-supervisor: Laura Slaughter, IDI

Norwegian University of Science and Technology
Department of Computer and Information Science

Abstract

Medical researchers today often work in a highly collaborative environment. The methods used to collaborate often include face to face meetings, unnecessary e-mail discussions, and huge amounts of documents sent and received in order to cooperate, share and contribute to each others work. The collaboration gets even more challenging as researchers today often do not work from the same location. For this study, the purpose has been to explore if there are ways in which the everyday work of dementia researchers can be supported and made more efficient by the use of web-based collaborative tools.

A case study including four researchers related to dementia research has been conducted. A series of interviews with the researchers were done, and based on the findings from these, a prototype has been suggested and evaluated by the participants. The main findings shows that there is a need for further support during collaborative and coordinating activities in order to make the research process more effective. A web-based collaborative tool made with the medicine 2.0 approach in mind is suggested, which includes social networking, supporting collaboration between different actors, and openness and sharing between these actors.

The main finding from this study is that there is a need for improved methods to support the collaboration and coordination between dementia researchers. Allowing the researchers to create social networks including sharing and collaboration related to files was seen as effective. It had potential of reducing the time spent revising documents as a result of tagging and notification functionality, in addition to improve coordination between colleagues because the awareness of what other project members were doing increased. The study also shows that there is a potential in using patient-produces sources such as blogs and illness communities in research, specially within qualitative research and social welfare studies. A need to improve the visualization of available data during the examination phase was discovered, however the need was not so strong as for improving collaborative activities.

Keywords: Dementia research, data visualization, Computer-Supported Collaborative Work, medicine 2.0, social networks.

Preface

The school year 2010/2011 was spent writing a master thesis within the field of medical informatics. It considers new ways in supporting dementia researchers during their everyday work. The completion of the thesis represents the end of five great years of studying Computer Science, with a specialization in System Development and Human-Computer Interaction at NTNU.

First of all, I would like to thank family, friends, and other people that may have crossed my road during these five years at NTNU. All of you have provided me with support, understanding and patience during the process. I appreciate any advices given, moral support, and discussions on the way.

The supervisors for the thesis have been Pieter Touissant and Laura Slaughter in the Department of Computer and Information Science (IDI). Laura has been my main contact, providing me with valuable feedback through the process. I really appreciate it.

Last, but certainly not least, a special thanks is directed towards the four participants in the study, who well willing gave of their time, knowledge and patience to an eager, blue-eyed master student. If it where not for them, this thesis would not have seen the light of day.

It is true as they say, writing a master thesis truly is a learning process. If learning is the goal, it has certainly been achieved.

Table of contents

List of Figures

List of Tables

1. INTRODUCTION

The subject of this study is to explore if there are ways in which the everyday work of dementia researchers can be supported and made more efficient by the use of different tools. The number of dementia patients is expected to increase during the next 40 years, so finding effective ways in doing research to prevent and treat dementia is important. This chapter starts by introducing the research area, followed by the specific research questions for this study. In the end, the expected result and an overview of the structure of the thesis is provided.

1.1 Research Area

Researchers often work in project groups, where some of the participants are seated in different locations. The collaboration among them is influencing the success and progress of their research. Some well-known methods for collaboration today are face-to-face meetings, discussions by e-mail, and telephone conferences. Also some Computer-Supported Cooperative Work (CSCW) systems that are intended to support the collaboration exist, but don’t seem to be in use to a great extent. A CSCW system could be adopted as an effective way of supporting the researchers during the research process. CSCW is argued to be more of a design principle, as a system designed for one specific user group might not work as intended in another user group. This may be some of the reason why it has not been adopted to such a great distinct by medical researchers yet.

In addition to the collaboration between researchers, another challenge in medicine is the great amounts of data available. Getting an overview of what data is available, and getting aware of the relationship between different variables, is a challenging yet important task. Effective ways in presenting such data is needed. In addition to the large amount of data, several quite different sources of research data exists. Still, the health professional-produced sources like the health record seem to be the preferred choice. There are already several methods to extract and use this information, for example in decision support systems that use data mining techniques to find patterns and relations between symptoms and diagnosis (Iavindrasana et al., 2009). Also patient-produced sources could provide valuable research data. Blogs provide a great deal of information that has previously been unavailable for researchers, and in other fields such as ALS (Amyotrophic Lateral Sclerosis), patient-written text has led to the discovery of a possible treatment (Frost and Massagli, 2008). The use of such data should be considered also in other medical research fields.

1.2 Motivation

This thesis considers the field of dementia research. Dementia is a disease affecting the older part of the population, and influences patients' cognitive abilities such as memory, language, and attention (Engedal, 2003). In Norway, there are about 60.000 patients suffering from dementia, and calculations anticipate a doubling within the next 40 years (Hjort and Waaler, 2010). This represents a large part of the population, and a huge load on elderly care - not to mention the effects on both patients and their caregivers in their daily life. More research is needed both in discovering, preventing, and treating dementia. Exploring ways in which this work can be made more effective will benefit the field of dementia research. The main scope of this thesis is to find ways to support their work more effectively.

Today, traditional methods such as e-mail and telephone conferences are heavily in use. This study aims at looking at more modern collaboration methods. The belief is that using methods from social networking, the field of CSCW and inspiration from the use of e-science tools, the work of dementia researchers could be made more efficient.

1.3 Research Questions

Getting to know how dementia researchers work is an important part of this Masters thesis. Finding ways to support their work by making collaboration and data sharing easier and more effective are key goals of the study. Next, the possibility for researchers to gather data from several different sources, including patients or relatives blogs, and different health registries in addition to the electronic health record should be exploited in order to expand the data material. What options are there, and how can the data be used? The specific research questions for this thesis are as follows:

> **RQ: Are there specific web-based, social collaboration and data visualization methods that can improve the efficiency of dementia researchers' work?**

As a part of the research question, several sub-questions should be addressed:

-How do dementia researchers conduct their studies, share data and work together during the research process?
-How do they use technology to support this process?
-Are there any tools that can be provided to increase the efficiency of dementia research, and how do these tools affect the research processs?
-Are there any additional data sources that could potentially be used within dementia research in addition to those used today?

1.4 Resulting Prototype

A goal of the research was to suggest a prototype of a tool aimed at supporting researchers in their everyday work. Development of the tools is based on findings from the study and evaluated by dementia researchers. Based on the results from this evaluation, recommendations are presented as for what functionality is seen as useful and should be further developed. In addition to the prototype, a discussion of today's practice throughout the dementia research process is discussed. The data collection and various data sources are a part of the study, as it represents an extensive part of the research process. Insight into the researchers' everyday work, and their needs and requirements for a supporting tool are key issues discussed.

1.5 Structure of the Report

The report is divided into several smaller parts, starting with a thorough introduction to the research field and the related theory in chapter 2. A case study of dementia researchers has been conducted, and a user-centered development approach has been used to develop the prototype. Both methods are further described in chapter 3. Chapter 4 gives a report on the results from the first round of interviews. Based on the findings from the interviews and other data collection, a prototype was developed which is also presented here. The prototype was presented to the researchers followed by an evaluation of the proposed functionality, and is presented in chapter 5. A discussion and analysis of the results from the study is presented in chapter 6. In the final chapter, chapter 7, a conclusion is made based on the research questions. The chapter also includes further work to the system, and interesting topics to examine closer are suggested.

2. THEORETICAL BACKGROUND

In this chapter, the theoretical background is presented. The chapter starts with an introduction to cooperative work, mentioning both Computer-Supported Cooperative Work, and the more recent field of medicine 2.0. In addition to cooperative work, getting an overview of the available data is a great part of the research process, and has been a part of this study. The information search process and visualization theory is summarized in 2.2 and 2.3, before the chapter close by presenting several existing visualization systems in section 2.4. Principles and lessons learned from all af the presented systems have been used to inform this work.

2.1 Cooperative work

2.1.1 Computer-Supported Cooperative Work

Computer-supported cooperative work (CSCW) was established as a term during a conference in 1984 (Grudin, 1994). Such systems are intended to help people work together more effectively. The human-human interaction is considered within the field, as CSCW systems are intended to support the collaboration between people working together. *"Computer-supported cooperative work is computer-assisted coordinated activity such as problem solving and communication carried out by a group of collaborating individuals"* (Baecker, 1994, p.1). CSCW is intended to support both the collaborative work between people, and the techniques used for collaboration. The technology may include communication mechanisms, support of group activities such as collaborative writing, and a shared workspace where information and joint files are shared. All of these may be useful for a dementia researcher. Since the CSCW approach appeared in 1984, several possibillites have emerged during more recent years. Social networking and different web-based tools have become easy available, and have contributed to the field of CSCW. Nowadays, it is easier to develop custom-made CSCW systems to fit in with individual fields, which creates new possibilites for the use of such systems.

Three key aspects of CSCW have been described by Pratt et al. (2004), namely incentive structures, understanding workflow, and incorporating awareness.

Incentive structures are about creating systems in such a way that it is properly adopted by the people in an organization. The potential users must be willing to spend time and energy to both learn and use a new system. Some benefits must come out of it, often in the form of more efficient work, higher quality of work, or other improvements compared to not using the system. Grudin (1988) argues that a common problem when adopting CSCW systems is that the persons doing the extra work are not the persons experiencing the benefits from the system. This means that one should make sure the researchers experience the benefits themselves - using the system must make their work feel more efficient or possibly easier in some way. Medical researchers are already under great time pressure, so the chance that they would adapt a CSCW system that doesn't seem to be supporting their own work is not that high. The benefits must come out clear in order for the system to be adopted.

Pratt et al. (2004) also mentions the issue of workflow, which relates to the activities done by each individual in order to successfully complete a work task. A CSCW system must consider the existing workflows in an organization before developing solutions to support it. Any change to the existing workflow in an organization must be justified by some perceived benefit in order to be accomplished. Grudin (1988) mentions that the design process often fails because the decision-makers ordering a system often see the benefits for themselves, without noticing that it leads to extra work by others. For medical researchers, this is a key issue. They have to be able to use whatever tools and methods they are used to, and familiar with. There are numerous tools and methods available throughout the research process, and researchers may use different approaches. Making a tool that allows individual workflows might prove to be a crucial point when considering adoption. To become aware of the existing workflows and decision making processes in an organization, a CSCW system developer must use more anthropological methods. This means longer lasting observation of groups, in-depth interviews, and other methods to get an insight into the social and cultural factors in the organization. This is some of the reason for the choice of data collection methods, which is further presented in the next chapter.

The last factor mentioned by Pratt et al. (2004) is awareness, which in this case relates to becoming aware of the activities done by others to provide a context of your own work. Collaborative work will be more efficient, because one is able to adapt activities according to the others. Medical researchers often collaborate with others, and are dependent on other researchers in order to get their own work done. Becoming aware of what colleagues are actually doing could lower this dependency some, because each individual researcher is provided with information about what the status of others are. Awareness could also be about knowing the status of a project - are they waiting for some permissions, do they have data available, and are the project approved by the management? Such questions could easily be answered if awareness were a supported feature in a collaborative system. Awareness has been a great part of the discussion around this study, as it allows each individual researcher to plan their work according to the others without too much effort of finding out the current status. The ability to make better plans and decisions without too much effort is believed to make the process of conducting research more effective, which is one of the main goals for this study. Easier communication and possibility to share the current status of individuals, documents, datasets, and entire projects has been considered during this work as it is considered valuable.

A CSCW system for biomedical research has already been proposed by Stark et al. (2008). The authors give a presentation of the main requirements they made for the system, many of which has also been adopted in this study. Among others, they mention flexible user and role management as an issue, as data protection and sharing are key issues. Access to data of such private character as medical data must be strictly handled; so that only the people allowed viewing certain data actually have the possibillity to view it. Further, they mention that shared files should be presented to the user as if they were located in the same physical location to hide complex data structures from the user. This makes the document retrieval and overview of the available documents much more intuitive, as the user doesn't need to consider additional, perhaps complex issues that are out of their control.

Stark et al. (2008) also mentions that there are several different kinds of users for such a system, so supporting flexible presentation of the data, as well as tools to handle the data becomes important. Both medical researchers, statisticians, and technicians might be a part of a research group, so a tool has to support different professional backgrounds in order to have the system adopted by the whole group. Finally, sharing knowledge and communication are key points. The authors mention that web 2.0 technology can assist in supporting the cooperative functions, which is a great part of a researchers work. Even though several key aspects to support collaboration is provided in their system, it seems to lack more extensive functionality to support awareness and coordination. Both of them are affecting the efficiency of each individual researcher, as some tasks are dependent on others, especially in a collaborative environment. A way to better support the collaboration and awareness should be provided.

Carole Goble talked about collaboration in the context of e-science, and it sums up a great part of this study quite well.

> *Real progress depends on pooling know-how and results. It depends on collaboration and making connections between ideas, people, and data. It depends on finding and interpreting results and knowledge generated by scientific colleagues you do not know and who do not know you, to be analysed in ways they did not anticipate, to generate new hypotheses to be pooled in their turn.*
>
> (Goble, 2005, p. 1)

Even though related to e-science, several of the aspects should be an important part of a tool that supports project groups where the participants are relatively well known to each other. As for this study, the main focus is smaller research groups where all the participants are familiar with each other - the scope is a bit narrower than that of e-science. A collaborative tool of smaller scope could anyway turn out to expand into an approach where researchers cooperate with unfamiliar persons as well. The sharing of knowledge, experience, data, and results to an even greater degree definitely has the potential of adding value to research.

The extensive sharing of data doesn't only provide benefits. Some challenges are crucial to overcome, especially those of information security when related to sensitive health data. However, some of these issues disappear when the data is anonymized. To allow such sharing of patient data that one prefers considering a CSCW approach, security measures must be taken to assure that only the person allowed to view the data gets access to it. Proper authorization, encryption, secure storage and several other techniques should be used to ensure the patient privacy and data security. It is considered as one of the most important aspects in CSCW systems for medical research, although it is not further discussed in this report. It has however been argued that when certain principles are carefully handled, storing such data in Internet databases are both effective and secure (Marshall and Haley, 2000). The sharing and communication on Internet allows for several new features that is better supported through a web-based site than running on a local network, as a web-based site easier allows access for researchers outside, for instance, the local office. The medicine 2.0 approach further investigates the possibilities that proceed using a web-based collaborative tool.

2.1.2 Medicine 2.0

Later years, new forms of collaboration have been introduced through web 2.0 technology. A final defintion of what medicine 2.0 actually is, seems yet to be missing. However, core concepts of medicine 2.0 has been mentioned as including web 2.0, patients, professionals, social networking, health information, collaboration, and change of health care (Van De Belt et al., 2010). One proposed definition is as following:

> *"Medicine 2.0" applications, services, and tools are defined as Web-based services for health care consumers, caregivers, patients, health professionals, and biomedical researchers, that use Web 2.0 technologies and/or semantic web and virtual reality approaches to enable and facilitate specifically 1) social networking, 2) participation, 3) apomediation, 4) openness, and 5) collaboration, within and between these user groups.*
>
> (Eysenbach, 2008, p.1)

The five points mentioned by Eysenbach are major aspects in medicine 2.0. He mentions the three main involved groups of medicine 2.0 as patients, health professionals, and biomedical researchers. The social networking between these groups opens up for new ways of performing medical research, as the patients themselves may serve as an information source. As already explained, the patients might contribute by giving researchers and physicians access to other perspectives of their disease through for instance blogs or forums. The new relationships between these groups create new possibilities for collaboration, which for instance can be seen in the PatientsLikeMe project, which are further introduced later. For this study, the main focus is on the relationship among researchers, although it does not exclude any possibilities to expand later work into including also the practitioners and patients. Creating and establishing a social network between researchers before including other user groups might provide an easier transition to more social, collaborating work methods. In time, engaging all three user groups has many potential benefits, as it opens for new ways of collaboration, data collection, sharing, and maybe also preventive care as the patients awareness themselves increases.

For the medicine 2.0 approach to be successful, participation is a crucial point; for instance, researchers must be willing to cooperate and commit to other researchers through taking a part in the formation of a social network, each giving some contributions to the other participants. Another key success factor for social networks is the issue of openness - researchers, patients, and health professionals must be willing to share data, experience, and knowledge with others. The sharing and openness related to information opens up for apomediation, which in this case means that the users help each other navigate through the ocean of information. A well-known example of this is Amazon[1], where users recommend items to other users.

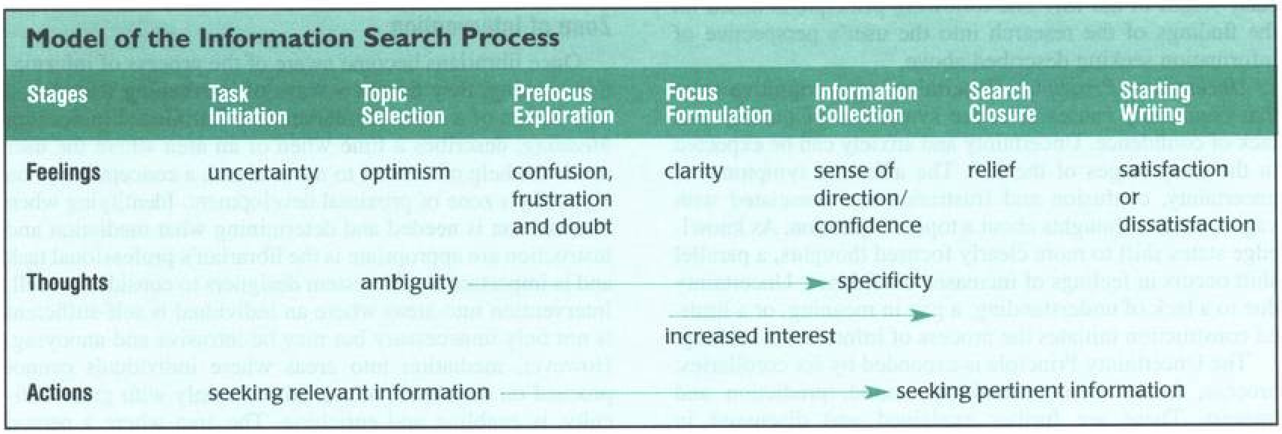
The use of social media within healthcare has already been investigated. Several illness blogs exists, being maintained by either patients or their relatives. The information provided through such blogs might have value at least for nursing research, which focus on quality of life, illness management, and handling symptoms (Heilferty, 2009). The content of medical blogs might serve as a valuable information source, and must be considered especially with the emergence of medicine 2.0 applications. Other sources than the traditional health-professional produced sources should be further investigated. In addition to blogs, wikis, online health profiles, and discussion forums also might provide some of the same information.

---

[1] http://www.amazon.com

2.2 The Information Search Process

In addition to collaborative activities, a main part of the researchers' tasks is of course the actual research process. According to Marchionini, *"Information is anything that can change a person's knowledge"* (1997, p. 5). A great part of the research process, is acquiring information and turn it into knowledge. During their work, several different information needs arises, for which the relevant data must be collected in some way. The process of gaining information has been described by Kuhlthau (1999) in the Information Search Process, as illustrated in figure 1. The model consists of seven phases, where a set of common feelings, thoughts and actions during each phase is stated.



Model of the Information Search Process

Stages	Task Initiation	Topic Selection	Prefocus Exploration	Focus Formulation	Information Collection	Search Closure	Starting Writing
Feelings	uncertainty	optimism	confusion, frustration and doubt	clarity	sense of direction/ confidence	relief	satisfaction or dissatisfaction
Thoughts		ambiguity →		→	→ specificity		
				increased interest →			
Actions	seeking relevant information →			→	→	→ seeking pertinent information	

Figure 1: The Information Search Process (Kuhlthau, 1999, p13)

Kuhlthau (1999) introduces the concept of uncertainty during information seeking, claiming that the uncertainty increases during the process. One could expect that the uncertainty decreases as more information is gathered, but research shows that a dip in confidence is experienced during some of the early stages. This is often due to conflicting and inconsistent information. As Kuhlthau explains, the exploration phase is often the most challenging, because of difficulties in expressing exactly what one is looking for. For a researcher, this implies that the available data should be visualized, in order to support the researcher in getting overview and assuring that the conclusions and results based on the data are correct. One way to do this is to give an interactive, self explanatory overview that is easy to grasp and manipulate as the comprehension increases.

A part of the information seeking process within the medical domain, might be to view and compare different medical records. Lifelines2 is a system intended to allow this comparison, and during the work with it, Wang et al. (2010) constructed a process model to do that, which also relates to Kuhlthaus model. Some additional phases are introduced by Wang et al., including the data acquisition and data examination phase. While working with sensitive information such as personal medical data, the data acquisition part takes time - not only to gain access, but also pre-processing the data, and making them anonymous. The step of examining the data is also an important step for all researchers. While doing this, a researcher assures himself that the gathered data seems valid, and also gain an overview of the data material available.

A well-known method to get an overview over data is by visualizing it. Presenting data visually often leads to the discovery of patterns that could be hard to discover otherwise (Shneiderman, 2002), for instance polarized data, or erroneous values that could bias a statistical analysis. During the examination stage, one could find previously undetected patterns in the data, or one might be able to formulate a hypothesis (Wang et al., 2008). However, proving a phenomenon often requires more thorough statistical analysis.

2.3 Information visualization

As explained, information visualization could potentially add value to the research process, as studies shows that it does provide support during several of the stages in the research process. Visualising the available research data has been a topic of discussion during this study, and should be further discussed.

Information visualization can be viewed as the process in which a human creates a mental model of data. Spence (2007) provides a model that shows the interaction with data in order to create a mental model, as shown in figure 2. It shows that raw data is presented to the user in an organized matter on a display, followed by a perception and interpretation of the information by the reader. Based on earlier experiences and mental models, a decision is made whether the information need is fulfilled.

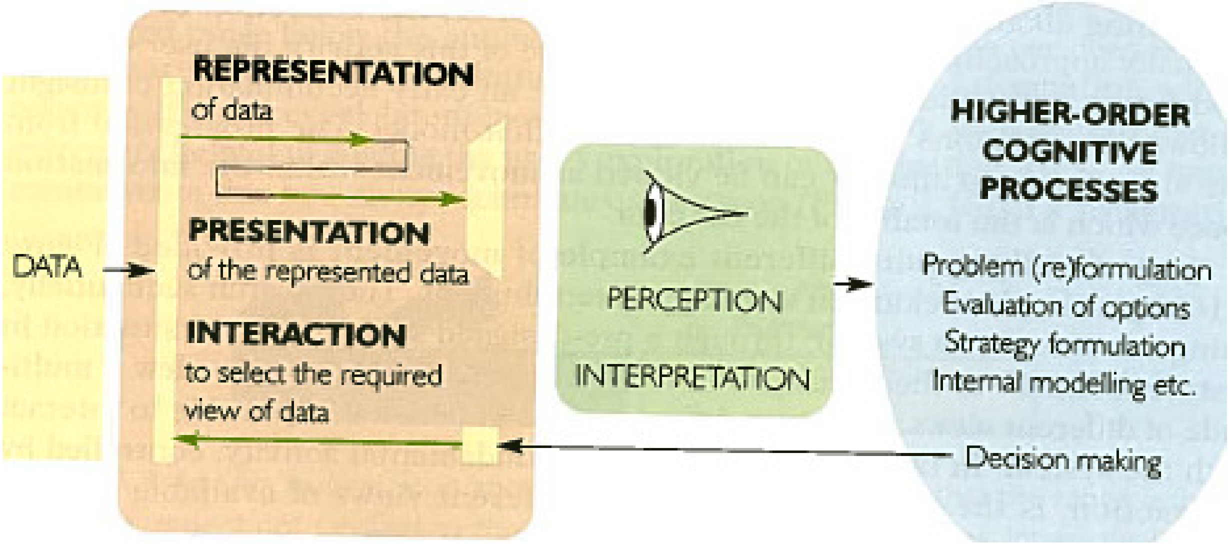


Figure 2: The Information Visualization process (Spence, 2007, p.26)

Spence (2007) states that information that is visualized, is often easier and faster to understand, and the mental images often lasts longer than raw data. These properties can be used to ease the cognitive load on the human mind when processing information. During research, the ability to gather and get an overview of huge amounts of data is important, which is why the use of information visualization tools for researchers should be further developed. Visualization is a human cognitive ability, and has originally nothing to do with computers (Spence, 2007). A computer can however assist in the forming of mental images and insight, among others because of their rapid access memory, computation speed, and possibilities to display data graphically.

An important implication to provide visual representations of data is to gain insight, and discover new relations between data. One successful history that does exactly this, is stated by Tufte (1997), where a doctor was able to discover and remove the source of a cholera epidemic in London by plotting the number of deaths from cholera on a map. By doing this, he discovered the relationship between a specific water pump and the outbreak of the epidemic. A part of the method to gain this insight was to view the data in an appropriate context, and in this way evaluates the relationship between cause and effect.

Different visualization techniques give a different view on, and context for, the data. Choosing the correct technique is important to gain insight and see new relationships. However, it can be difficult to know in advance what visualization technique will prove to be the most valuable, so being able to switch between different views and visualization techniques in a research tool can be very useful. Another reason why several views should be supported is that a research group often consists of participants that work by different methods, or with a different point of view. Also, the character of the data might be very distinct, as medical research makes use of both qualitative and quantitative data. As explained earlier, different data sources are considered such as blogs and online patient communities, and such data should also be viewable through a cooperative research tool. Even though the potential users, data sources and methods all should be supported in such a tool, they have some tasks in common. Schneiderman (1996) suggests seven main tasks, provided in table 1.

Overview	Get an overview of the collection at hand
Zoom	View closer the items of interest
Filter	Remove uninteresting items
Details-on-demand	View details of a selection when necessary
Relate	See the relationships between items
History	Remember the user history to enable undo, redo, etc.
Extract	Allow saving of sets of items with the purpose to extract later

Table 1: Seven information visualization tasks (Shneiderman, 1996).

A visualisation system should support all these tasks as a minimum. Wang et al. (2008) argues that also align, rank and filtering should be included as these are frequently used operators when viewing medical records.

Even though several possible views and tasks should be supported, one should avoid too much zooming and panning because the constant interaction and changes disturbs the researchers visual memory, which makes it harder to discover patterns and compare several records (Wang et al., 2008). A consequence of this is that huge amounts of data have to be presented on one screen. Both horizontal and vertical overviews prove to be valuable, but increasing the information density is still an important issue. Research shows that in the visualization system Lifelines2, which is a system to visualise and compare multiple medical records, *every* session contained scroll-activities (Wang et al., 2010). It should be a goal to lower this number, while still offering the same amount of data.

2.3 Existing research tools

Several information visualization tools aimed at medical research already exist, and are a current topic of research. Even though existing research tools make secondary use of health data to support the researchers work, they focus mainly on the information visualization part of the research and not so much on collaboration. Secondary use of health data is when patient health data is used for other purposes than originally gathered for, and include non-direct care like research, analysis and quality measurements, among others (Safran et al., 2007). The reuse of data could lead to a more efficient cycle of data collection, research, and knowledge generation. This is especially helpful if combined with clinical decision support systems, which gives advice to the clinician based on results from previously treated patients. In the future, such cooperation between stakeholders should be further investigated. Medicine 2.0 might be a great solution to support this knowledge- and experience sharing between different actors.

Both information visualization and secondary use of health data has great potential value. However, there are still several issues to overcome. Gaining access to the records has political, ethical, and technical challenges. How can one assure that the patient allows reuse of their record? Do they have to give permission to every entry, or can a global permission be given? Even though these topics are not discussed further in this study, they are considered as key issues to resolve before an actual research tool to support secondary use of health data can be developed.

For many researchers, gaining overview over large collections of data material is a central task. Earlier attempts to do this were often through statistical analysis, which provided compact, numerical values or tables. They are often easy to understand, but lack the ability to show distributions and outliers. Some relationships between numbers can also be hard to discover, for instance, using a linear model when the relationship is really exponential. Visually presenting the data set can be a much more intuitive way to find the correct relationships.

A suggestion to combine the statistical data mining with the visual presentation has been made when inventing discovery tools (Shneiderman, 2002). The two approaches stem from different research methodologies. First, there is the reductionist which favour making hypothesis in advance, changing only one variable at a time during controlled experiments. Critics claim that the approach often leads to a narrow sight, where the researcher might miss certain phenomena in their eagerness to prove or disclaim the hypothesis. Second, there are researchers with a more exploratory point of view, which gather large amounts of data, and then start searching for a pattern. However, the generality and validity of these results can be questioned. Shneiderman argues that combining these approaches in one tool might give the best from to worlds. One can do this by allowing the user to explore the data visually, and then apply statistical tests when the researcher suspects a relationship between certain data. The visualization of the data can often help in choosing the correct statistical model, leading to faster results. However, the computers ability to do complex analysis on large datasets must not undermine the user. Their feeling of control is extremely important in order to create a valuable tool in which one can trust.

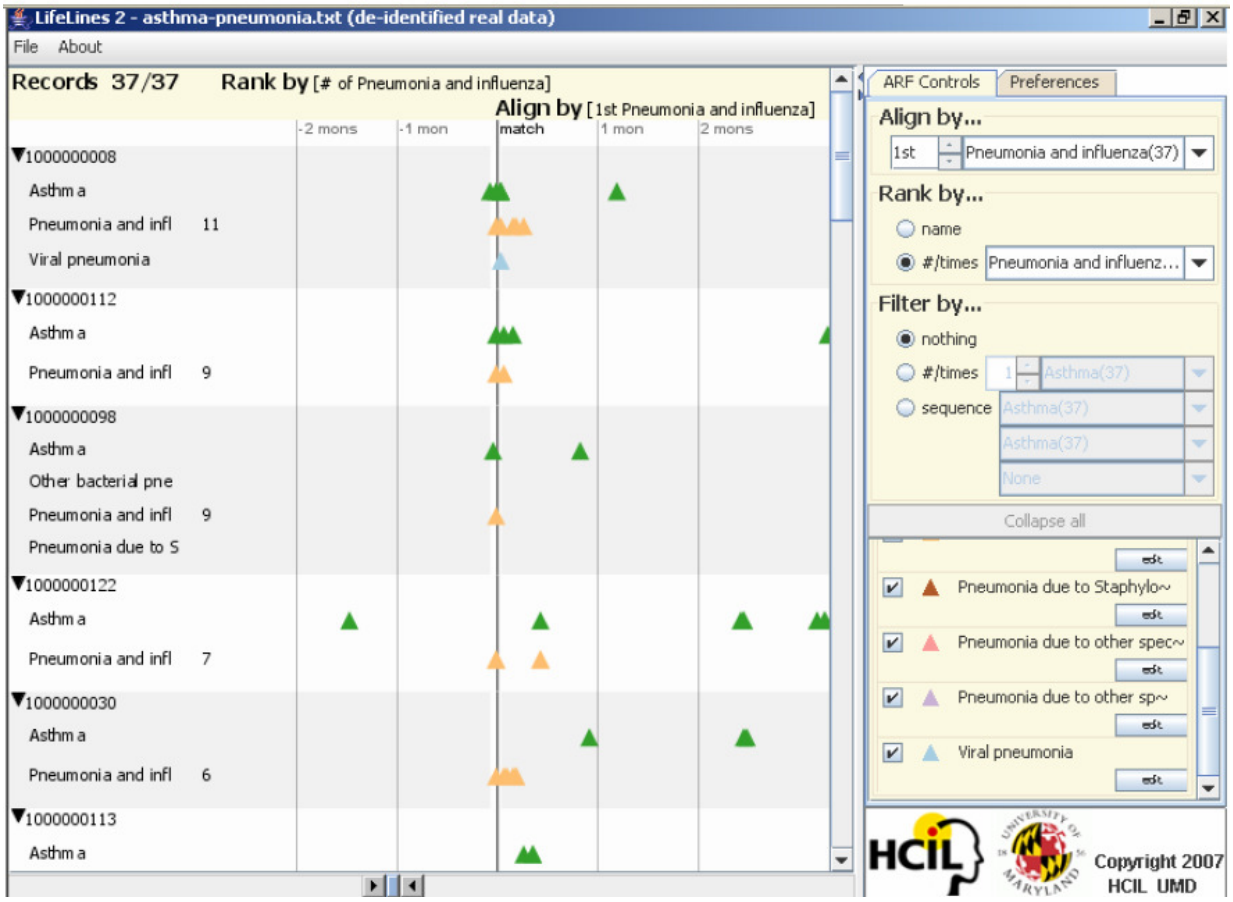
Some visualization systems that already make secondary use of health data have been suggested, and are presented below. The presented systems have all been an inspiration for the suggested prototype in this study with their lessons learned and topics discussed.

2.3.1 Lifelines

Lifelines is a tool developed for clinicians to represent an entire patient record, with a focus on summarizing and giving an overview using timelines (Plaisant et al., 1998). It has been developed with the information seeking mantra in mind (Shneiderman, 1996) ; *Overview first, zoom and filter, then details-on-demand.* Lifelines has in later years appeared in an extended version, Lifelines2, which includes support for viewing multiple records at once (Wang et al., 2008). This extension makes Lifelines a more suitable product for researchers, who often need to compare and get an overview over several records. One technique that has proven successful in comparing several records is by aligning events on a timeline by some sentinel event, for example the occurrence of an admission. Such alignment can lead to the discovery of previously undetected patterns.

The align-rank-filtering (ARF) framework that they use heavily prevents too much zooming and panning. By using the three operators, the user is allowed to define what data they want, in what order, and how it should be presented. It also decreases the need to remember too many details, which leads to an ease on the human cognitive load when comparing records.

Researchers typically work with a lot of data at once, so decreasing the cognitive load would contribute to an easier work situation for the researchers. However, it is argued that to be useful the ARF-framework must be data-driven, and easily changed in a dynamic matter - or else, the function will not be used to the extent it deserves. The creators experienced an improvement in discovering relationships between events when using the ARF-framework. A screenshot of Lifelines is showed in figure 3.


LifeLines 2 - asthma-pneumonia.txt (de-identified real data)

Records 37/37

Rank by [# of Pneumonia and influenza]

Align by [1st Pneumonia and influenza]

Align by... 1st Pneumonia and influenza(37)

Rank by... ○ name ● #/times Pneumonia and influenz...

Filter by... ● nothing ○ #/times 1 Asthma(37) ○ sequence Asthma(37) Asthma(37) None

☑ Pneumonia due to Staphylo~

☑ Pneumonia due to other spec~

☑ Pneumonia due to other sp~

☑ Viral pneumonia

Figure 3: The LifeLines2 interface (Wang et al., 2008, p.4).

2.3.2 LifeFlow

LifeFlow is a visualization system that focuses on giving an overview over event sequences (Wongsuphasawat et al., 2011). The tool is developed as a standalone tool, but can also be integrated with Lifelines2. It provides an effective way to give an overview over the frequency of event sequences, and the time slot between different events. LifeFlow is scalable, and can present a large number of records without requiring any additional screen space. This is done by showing the frequency of event sequences relative to each other - the higher percentage of a specific event sequence, the larger space on the screen it takes. As one can see in figure 4, the most common events end up with the tallest event bars. Each bar represents a specific event, such as arrival to the hospital, entering the Intensive Care Unit, etc. The whitespace between different events symbolizes the average time between the events.

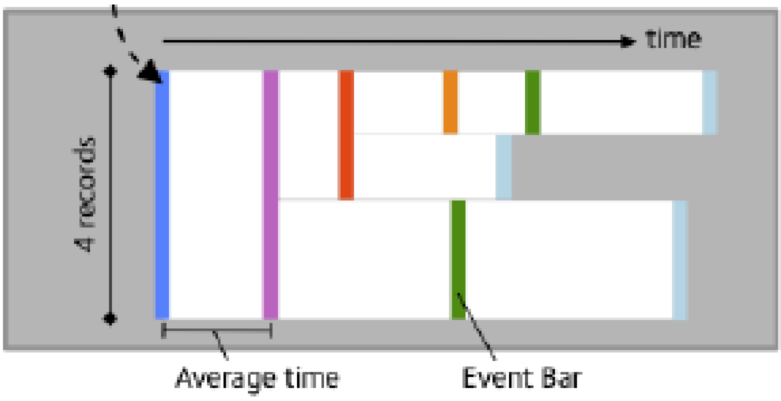
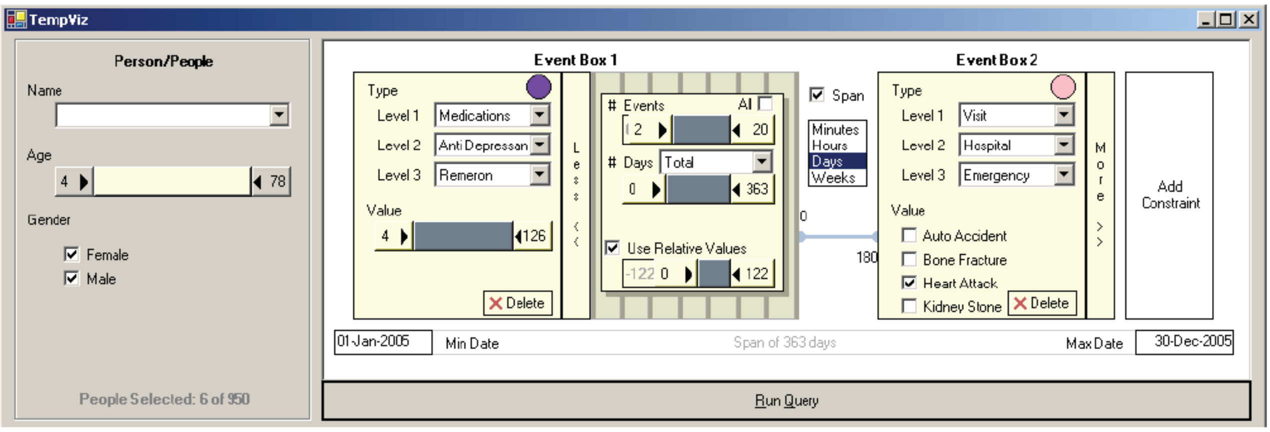


Figure 4: Visualization method from LifeFlow (Wongsuphasawat et al. 2011 , p.2)

This visualization technique is effective in giving overviews, but researchers also need more accurate data and details. This can, for instance, be provided by letting the researcher choose one or more event sequences to view in further detail. The authors have suggested using Lifelines2 to give a more detailed view, as LifeFlow is developed with the integration in mind, for example by allowing the researcher to zoom further in on certain patients.

2.3.3 PatternFinder

PatternFinder is an interface which focuses on visualizing queries over event sequences and their results (Fails et al., 2006). PatternFinder allows the user to query a dataset using a visual interface, instead of the traditional SQL-queries to find patterns and relationships between events. The query interface consists mainly of two parts; the first for filtering out specific patient groups, and the second for adding events and the temporal relationship between them (Fig. 5). The graphical way of specifying queries might appear as more appealing to the general medical researcher, as their speciality is not SQL. An interface like this provides a more intuitive way of specifying what they are looking for, without requiring knowledge about the structure of the underlying data.



Name:
Age: 4 – 78
Gender: ☑ Female ☑ Male
People Selected: 6 of 950

Event Box 1
Type
Level 1: Medications
Level 2: Anti Depressan
Level 3: Remeron
Value: 4 – 126

\# Events: 2 – 20 Al ☐
\# Days: Total 0 – 363
☑ Use Relative Values -122 0 – 122

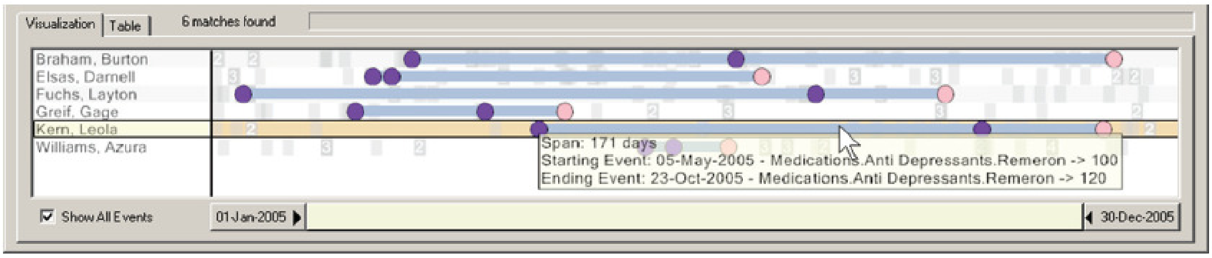
☑ Span Days

Event Box 2
Type
Level 1: Visit
Level 2: Hospital
Level 3: Emergency
Value
☐ Auto Accident
☐ Bone Fracture
☑ Heart Attack
☐ Kidney Stone

01-Jan-2005 Min Date
Span of 363 days
Max Date 30-Dec-2005

Figure 5: Pattern Query Panel (Fails et al., 2006, p169).

The result of the query is also presented graphically (fig. 6). As seen below, there is shown one row for each record that matches the query. This makes the interface less scalable, and a solution that scales better to large datasets might be preferred by the users, because some of them needs to view larger data sets. Another disadvantage about this way of presenting the results is that it might be difficult to get an overview of the time span between the events. One solution to this might be to use an approach like in LifeFlow to present the relationship between the events. However, in LifeFlow only the average time span is presented, which might lead to wrong conclusions if the researchers don’t investigate the individual records further.



6 matches found

Span: 171 days
Starting Event: 05-May-2005 - Medications.Anti Depressants.Remeron -> 100
Ending Event: 23-Oct-2005 - Medications.Anti Depressants.Remeron -> 120

Figure 6: Pattern Result Panel (Fails et al., 2006, p170).

Even though there are some disadvantages, the query panel has been brought on, and the result panel has been used as an inspiration for showing groups of patients. A method to keep the overviews of time spans between events and still keeping the scalability, might be to sort the records by increasing time spans and then show the records for instance in an area graph. The user should still be able to sort and align the events as needed, to make the chances of discovering a pattern higher.

2.3.4 Knave II/VISITORS

Knave II is a tool for visualization and exploration of clinical data, with a specific focus on time-oriented data (Bar et al., 2004). They claim other tools do not support abstractions of domain-specific knowledge to the same extent as their tool. The creators stress the importance of showing not only the raw data, but also the aggregated abstractions. Knowledge is captured in an ontology that is shown in the main window, making clear to the user how different concepts relate to each other. This way, both sibling, parent, and child concepts seems clear to the user, making it easy to switch between both raw and abstracted data. There are also several different zooming tools in order to provide an overview of the temporal aspect, as shown in figure 7. The Knave II tool makes use of more statistical views than some of the other tools presented. Descriptive statistics such as the mean and standard deviation is available, and the default for the abstractions is the distribution of the duration of the values (Shahar et al., 2003). However, it supports the visualization of only one record at a time. It may therefore be more suitable for a clinician, for whom the system was originally designed. They already focus on combining data and knowledge, which fits in with the idea of clinical decision support systems.

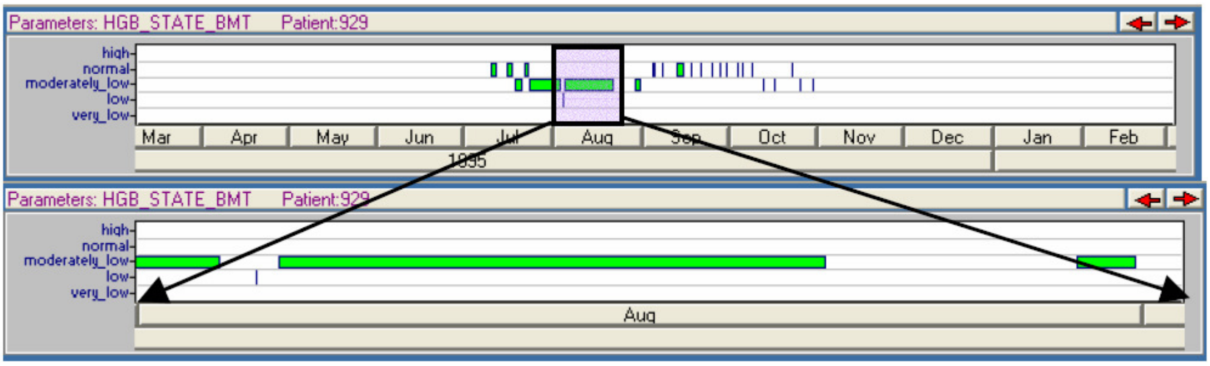
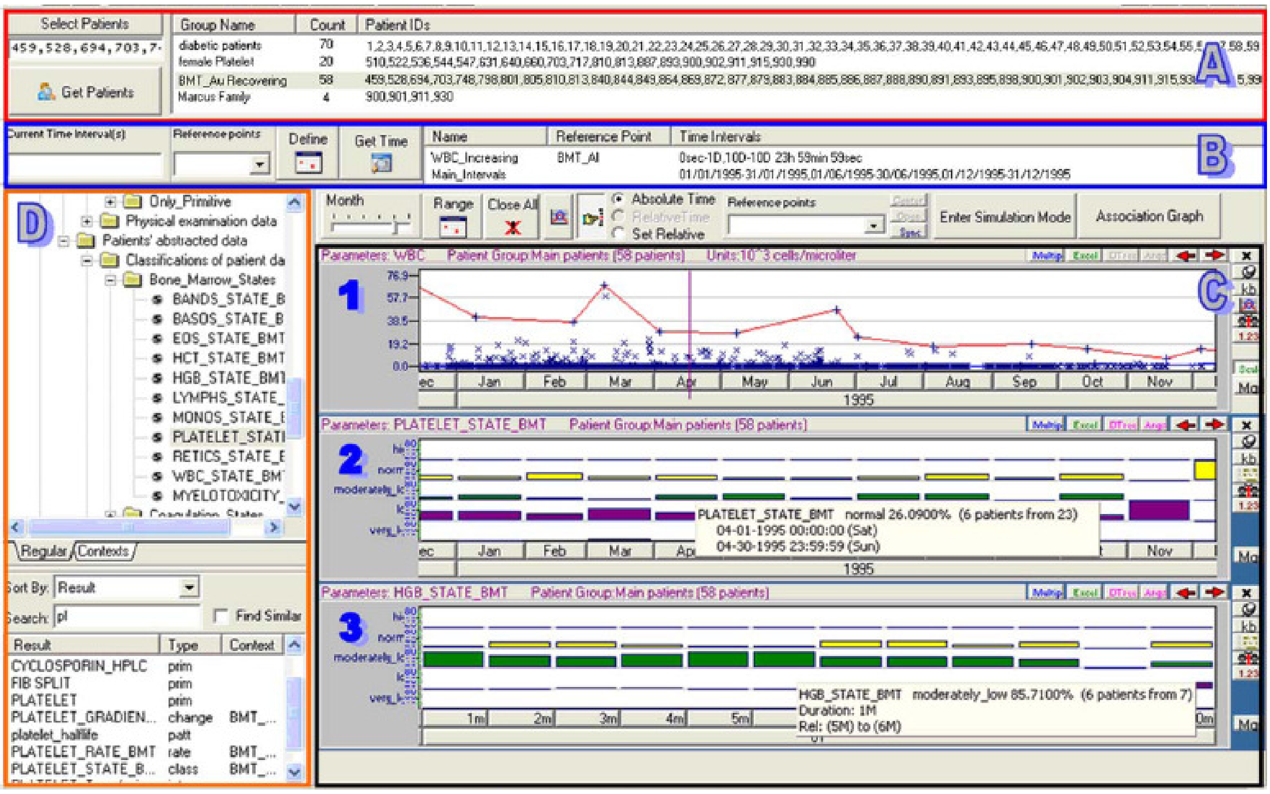


Figure 7: Bar et al. (2004) p.173

VISITORS (Visualization of time-oriented records) is an extension of Knave II that allows the user to view groups of patients (Klimov et al., 2010).



Group Name	Count	Patient IDs
diabetic patients	70	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,5 7,58,59
female Platelet	20	510,522,536,544,547,631,640,660,703,717,810,813,887,893,900,902,911,915,930,990
BMT_Au Recovering	58	459,528,694,703,748,798,801,805,810,813,840,844,849,864,869,872,877,879,883,884,885,886,887,888,890,891,893,895,898,900,901,902,903,904,911,915,93 5,99
Marcus Family	4	900,901,911,930

Name	Reference Point	Time Intervals
WBC_Increasing	BMT_Al	0sec-1D,10D-10D 23h 59min 59sec
Main_Intervals		01/01/1995-31/01/1995,01/06/1995-30/06/1995,01/12/1995-31/12/1995

Result	Type	Context
CYCLOSPORIN_HPLC	prim	
FIB SPLIT	prim	
PLATELET	prim	
PLATELET_GRADIEN...	change	BMT_...
platelet_halflife	patt	
PLATELET_RATE_BMT	rate	BMT_...
PLATELET_STATE_B...	class	BMT_...

PLATELET_STATE_BMT normal 26.0900% (6 patients from 23)
04-01-1995 00:00:00 (Sat)
04-30-1995 23:59:59 (Sun)

HGB_STATE_BMT moderately_low 85.7100% (6 patients from 7)
Duration: 1M
Rel: (5M) to (6M)

Figure 8: The VISITORS interface (Klimov et al., 2010, p.12).

Above, an overview of the interface is given (fig 8). It consists mainly of four parts; the part denoted by A contains different groups of patients that one can choose from, B contains different lists of time intervals, and part C shows graphs containing the relevant data (for example distribution of HGB state values over a month) for the chosen group of patients. Part D shows the knowledge-based browser over the domains ontology, which is one of the features that separate VISITORS from other similar products. VISITORS might better suit a researcher than KNAVE II, as it presents a group of patients in one screen.

2.3.5 PatientsLikeMe

PatientsLikeMe is a social, online community where patients share information about their medical history (Frost and Massagli, 2008), and is probably the system that is most unlike the others because the information is provided by the patients themselves. A screenshot from PatientsLikeMe is provided in figure 9. The system matches people with similar diseases, so that they can compare their development with each other. Patients comment on each others profile, communicate in private by sending personal messages, or simply view others profile. Such personal health records makes the patient an active part of their medical treatment, as is suggested by the medicine 2.0 approach. . Not only do they learn about their own disease, but also they might discover new treatments that have worked for other patients but have not yet become a part of the standard treatments offered. Patients give each other advice, recommendations, and support while coping with illness-related difficulties through this social network. Many of the patients using these communities suffer from severe and possibly incurable diseases. Talking to persons in the same situation is often experienced as comforting; and in addition the data they provide has proved to be useful for researchers. By providing benefits for both the patients and the researchers, the change that such a system will be adopted is even greater.

Information provided by the users of the system can be extracted and aggregated to create summaries for the entire communities, providing valuable statistical data. For rare diseases, such information might prove to be very valuable because of the limited number of patients suffering from the disease. However, the quality of the information must be evaluated. Also, one has to consider that patients edit their own records, telling the version from a subjective point of view, and might leave out important information. The quality and completeness can therefore not be guaranteed. Frost and Massagli's report (2008) shows that people with few options, for example those suffering from life threatening diseases, don't wait for science to finish lengthy randomized controlled trials (RCT)- instead, they drive research themselves. This urge to feed in data, keeping their profiles updated at all times, is definitely beneficial for researchers.

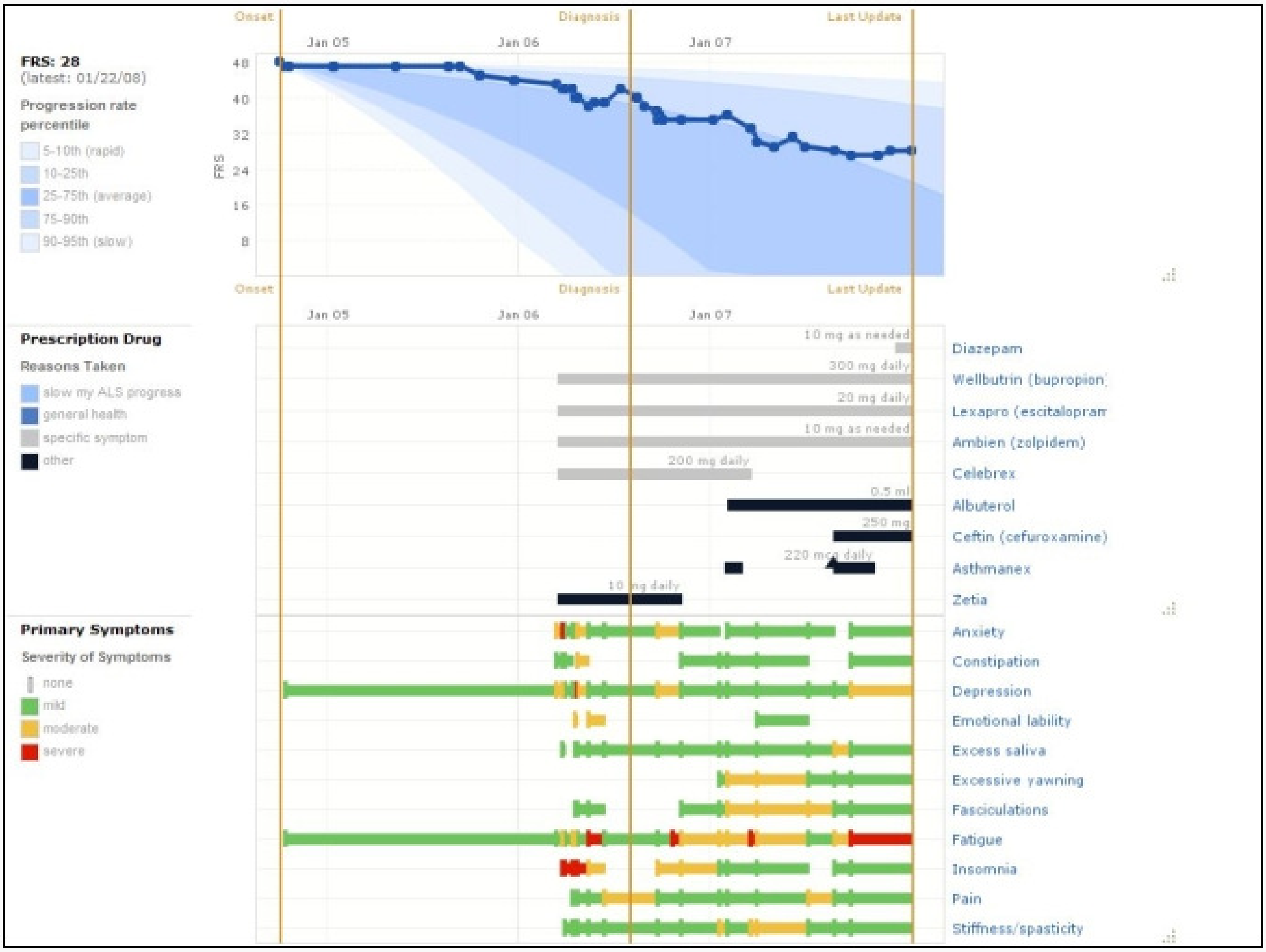


Figure 9: The PatientsLikeMe interface (Frost and Massagli, 2008).

3. RESEARCH METHOD

For this thesis, a case study considering dementia research has been conducted. This chapter provides an overview of the data collection methods used, as well as the prototype development method, which has been User-Centered Design (UCD). The final sections of this chapter discuss the expected results, in addition to some limitations both regarding the scope of the project and the methods used.

3.1 Choice of Method

Cornford and Smithson (2006, p.10) write that "*Information systems is about understanding what is or might be done with technical systems, how they come to be developed or chosen, and most importantly their effects and consequences in the world*". In this thesis, understanding how a web-based collaborative tool can support a researcher, and how such a tool might affect their work is crucial for its success. To achieve those goals, a structured research method is needed. Research methods are generally divided into a qualitative or a quantitative approach. Which method to choose, depends on what type of result one wants. Quantitative research was originally used within the natural sciences to study certain phenomena, and the results were often presented numerically. Qualitative research was originally developed in social sciences, to explain human behavior, and their social and cultural context (Avison and Myers, 2002). In this case, understanding the researchers and the context they work in is important to be able to develop useful systems, hence a qualitative approach was chosen.

Even within the qualitative approach, there are still several different methods - ethnographic, action research, and grounded theory - to mention a few. For this thesis, there are a limited number of participants available, a well-delimited research area, and focus on the use of an information system in a real-life context. For these reasons, a case study was considered to be the best suited. The case study method is widely used within the field of information systems. Avison and Myers (2002), mention three reasons for choosing a case study, which matches the scope of this thesis quite well:

1) The researcher can study the use of information systems in a natural setting, and this way produce theories from practice
2) The method helps the researcher to understand underlying processes, answering "how" and "why".
3) Case studies are an effective way to research areas where little work has been done previously.

In this thesis, a prototype was made for a narrow group of researchers. Most of the results from this thesis can probably transfer to other clinical areas. However, the focus here is on a defined group of researchers, namely dementia researchers. This choice was made to limit the task into a concrete group of researchers, hopefully with some similar challenges during their everyday work. The most used data collection methods within a case study are interviews, observations, and document analysis. Due to limited access to researchers within the area, the main data collection method has been a series of interviews, in addition to extensive literature review already presented in chapter 2.

3.2 Access to research area

Four researchers agreed to participate in the study. Two of them work as dementia researchers, and one is a chief psychiatrist who has also done work related to dementia research and dementia registers. The last participant works at HUNT (The Nord-Trøndelag Health Study[2]), which provides a health register for about 126 000 inhabitants in Nord Trøndelag county. The participants are all referred to as "the researchers" in this study, as they all can relate to research and work within either dementia or psychiatry.

Access to the domain is not so easy to get. The information one could get a hold of are in many cases sensitive, as they are related to personal, medical data. Access to the patients'

---

[2] http://www.ntnu.no/hunt/

journals or gathered data was not an option, as these are considered too sensitive. Even though they could have provided valuable insight into the research process - especially related to the data collection and idea creation phases - the patients privacy must be respected at all times. Even though the collected data was not an option to study, some document reviewing was done. The thesis considers other data sources, such as patient health writings that are publicly available. Also a topic list from annotated records and an overview of the content in health registers were considered. The document review is further presented below.

Because of the privacy issues, observation was a method that was found difficult to carry out. Also, the researchers work around tight schedules, so hours of observation might be too much to ask at such an early stage of the development process. Participants were recruited through e-mail or telephone, and found through existing dementia research projects, recommendations from people within the domain, or forwarded emails. Several researchers were asked to participate in the study, but only four were able to spare the time. On the basis of limited access to dementia researchers, and respect to the privacy issues, interviews were chosen as the main data collection method for this thesis. A total of six interviews were conducted; four previous to the creation of a prototype, and two interviews afterwards.

3.3 Interviews

Interview as a data collection method is a qualitative method where the researcher has a more or less structured discussion with the interviewees. Interviews can be divided into three different categories depending on the degree of structure (Robson, 2002). The semi-structured interview was best-suited to the research questions asked in this thesis. In a semi-structured interview, the researchers prepare some questions in advance, but are also free to ask questions in addition to those in the interview guide. Often, researchers do structured interviews to make sure the same questions are asked to different people. It is also easier to compare the answers for several interviewees. However, in this thesis, the number of participants was low, so there was no need to compare, quantify, or try to make a trend out of the answers given. Unfamiliarity with the domain increases the probability that several questions not thought of in advance come up during the interviews. The semi-structured interview allowed the flexibility needed to carry out this study.

For this study, the interviews were carried out in two rounds. First, all the researchers were interviewed in order to get an insight into their working methods and the dementia research process. The interviews were divided into three parts. The first part addressed the research process of each of the researchers. How do they gather information, in what ways do they collaborate with colleagues, and what tools do they use to get an overview of the collected data? The second part focused on different data sources. The patients' record is an obvious data source, but is there any potential value in using other sources, such as blogs and forum posts? And if so, for what? The last part concentrated on different research tools. Both the tools they use today and some of the research tools presented in chapter 2 were discussed. The main focus was on what tasks such a tool should support.

After the first round of interviews, the information collected was analyzed. Based on the result and the topics discussed, some requirements were formulated, and a prototype was developed. The second round of interviews centered on evaluation and user feedback on the prototype, as well as some follow-up questions from the previous interview. Two of the participants from the first round of interviews were visited. The interviews started with an introduction of the topics, followed by a 10-minute demonstration of the prototype. Some comments on the prototype were given during the demonstration, but most of the discussion followed after the presentation. The concept in general was discussed, together with specific functions and possible extensions and changes. . Interesting ideas and needs that did not emerge during the first round of interviews were covered in the second interview. The results from the interviews are presented and discussed in chapter 4 and 5.

The role a researcher takes with the interview participants is crucial, and can influence the result of the research. The contact with the participants is primarily through the interviews, so the appearance is important. Myers and Newman (2007) mentions that one of the problems during an interview is that the interviewee doesn't trust the interviewer, which can lead the interviewee to hold back information. An attempt to avoid this is making a good first impression - polite and trustworthy, making clear the intention to help the interviewee. This can be done through a well-prepared introduction before the interview, and also by asking questions that don't seem intruding. In this study, this was done by meeting the participants either in their office or a meeting room at their workplace. In advance, they got an email to briefly describe the thesis, and listed what themes were going to be covered in the interview.

This prepared the participants for what was going to happen, and hopefully led to an interview situation that was considered safe and comfortable by the participants.

The interviews started with a repetition of the thesis topic presented in the email, and some basic start-up questions that asked for the interviewee to present themselves and their work. Before the interview, the participants were asked if the interview could be recorded, in order to catch all the details. One of the participants was not comfortable with the recorder, so it was not used. This may have led to some missing information or points in that specific interview. However, notes were taken during all interviews to gather data, and in addition the other interviews were recorded and transcribed. Notes were considered useful especially in the cases where specific systems or images were shown, because the participant pointed, and in some cases drew, on the images. The results from the interviews were grouped by topics discussed, and similar opinions were merged. The topics discussed and the different opinions given are all discussed later.

The interviewers role during the interview affects the answers given (Oates, 2006). What level of experience within the domain the interviewer seems to have, could influence the level of detail provided by the informant. An important task is gaining trust, so the informants actually do give away the valuable information they poses. A part of this is convincing them about the intention to help. Also, the interviewer interprets participants´ responses given during the interview, which is important to be aware of. One should arrive without prejudgement or assumptions about the way they work. This way the interviewer can be open to all the information given. The semi-structured interview also opens up the possibility of asking follow-up questions, which are often considered the most important questions in retrospect.

3.4 Document review

For this thesis, four main document sources have been considered. First, a document presenting the content of a health record was reviewed. Along with personal experience within the health domain, it provided basic insight into what is actually in there. Second, a specific patient blog was considered. Several patients (or relatives) write blogs about their disease, including dementia patients. A specific blog was considered in this study, and an extract from this blog was shown to the researchers during the first round of interviews when considering different data sources.

What is important to be aware of when reviewing documents, is that they represent only a partial view, and express only what the authors wants them to express (Oates, 2006). This is especially relevant when reading patients blogs. The impression from reading those could be much distorted, and not be representative for the patient discussed. However, they do provide information about a patient that maybe isn't presented elsewhere, and could provide a new source of information that is valuable for researchers.

The last source considered, was information about available data in health registers. Several health registers exist, and an overview of what they contain are available online. They do not present the actual data values, but they give an overview of variables, number of patients, and in some cases a specification of collection method or the distribution of data values. Last, a document representing the data collection for a dementia register was provided by one of the participants in the study. This was a version of the actual scheme used during the collection to such a register. There were no plans to do content analysis on these documents, but the intention was to use them as valuable background material in order to get an insight into dementia researchers work.

3.5 Development Method

When designing a tool for dementia researchers, the focus is on the users and their needs in the specific case of dementia research. The choice of system development method for the proposed prototype has to consider this strong emphasis on the user. Several different system development methods exist, but the choice was to use the user-centered design (UCD) approach. The method ensures that new tools "*incorporate the needs, wants, skills, and preferences of the user throughout the tools' development*" (Morales Rodriguez et al., 2007, p.1). These are important goals when developing systems that are to be used within the medical domain.

The human-centered design process have been described in ISO standard 13407, and have been defined as an approach that focuses on making systems usable (ISO, 1999). The ISO standard claims that involvement of users during the process provides valuable insight to the context of use, the users' tasks, and how the users will use the system in the future. The early inclusion of the intended user, for example by the use of mock-ups, prototypes, and early user testing makes it easier to build the correct system - a system that works as expected, fulfills the users needs, and last but certainly not least, is accepted and used by the user (Morales Rodriguez et al., 2007). Such topics are considered very relevant in this work. A solid base for further exploration of the domain is important, so early inclusion of the users have been emphasised.

The ISO 13407 standard includes a description of the UCD. An adaption is shown in Figure 10, which represents the development cycle followed in this study. It starts by selecting the appropriate design method - in this case, the UCD. The model is iterative, and goes through 4 stages; identifying user needs, specifying requirements, development, and evaluation. After the evaluation, the cycle ends if the product satisfies the requirements. Otherwise, the cycle goes one more round, this way leading to an iterative development of the product. The product presented in this thesis went through one and a half round of the model, ending with an evaluation of the prototype followed by an interview to specify new requirements. As a result of this study and the experiences from the prototype evaluation, several recommendations as for what a collaborative research tool should provide are made.

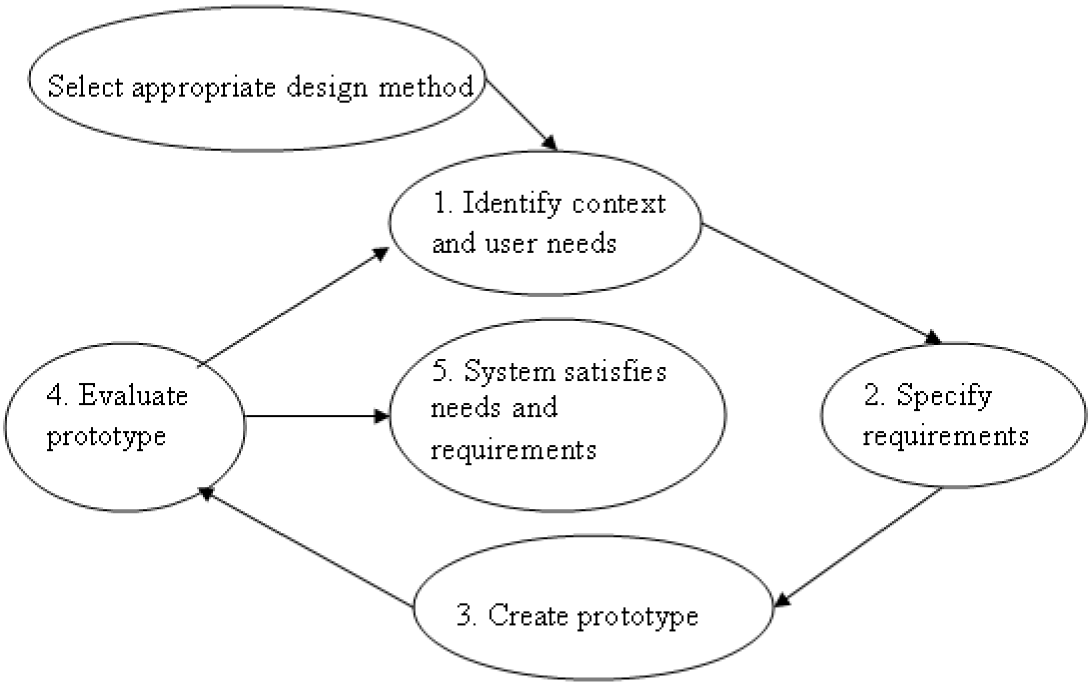


Figure 10: The Human-centered design method. Adapted from ISO 13407 (ISO, 1999).

> *In iterative design approaches, feedback from users becomes a critical source of information. Iteration, when combined with active user involvement, provides an effective means of minimizing the risk that a system does not meet user and organizational requirements*
>
> (ISO, 1999, p. 3).

The prototype was used to gather further information, and as a direct result of the proposed prototype, several new ideas and needs was discovered that probably wouldn't have come up using only regular interviews. Identifying needs at an early stage is beneficial both regarding time and economics in development projects. For this study, the prototype has been used as a mock-up to discuss ideas and possible functions. The experience from the first interviews, were that coming up with those ideas without a mock-up to look at was hard. Needs that were expressed by the researchers during the first round of interviews changed during the second round when they were introduced to the prorotype, so using a prototype as a way of gathering requirements was a good idea.

3.6 Limitations

A part of the result for this thesis will be a tailor-made prototype, which means that the suggested product is developed with a quite specific user-group in mind. The number of participants in this project has been a limited number of medical researchers within dementia, and their view cannot be seen as an answer for the entire research community. However, it serves as an example of one way to solve the needs and requirements from a medical researcher. Other researchers might have other needs, which should be further investigated in later work. Due to limited access to dementia researchers, observation has not been conducted. In a system that will be used in such a social and collaborative context, observation should be conducted in order to get an appropriate insight into the processes involved. However, the observations can also be conducted at a later stage.

The prototype presented in the study has been low-fi, which has both advantages and disadvantages. The greatest advantage is that it can easily be further developed in an iterative manner. Also, when it is obvious for the interviewees that the product is in an early phase of development, it often feels easier to come with feedback regarding the drawbacks of the product. The threshold of suggesting changes is lowered, and the cost of changing the prototype is minimal at the early stages. A clear disadvantage of testing with low-fi prototypes is that the user loses the sense of interaction and dynamics within the system. This makes it harder for a test person to get a correct impression of the system, leading to misunderstandings. Also, several key properties such as speed and the visual presentation are often hard to create in such a prototype.

4. RESULTS INTERVIEWS ROUND 1

The interviews unveiled several needs and possible points of improvement regarding the support researchers get while going through the research process. This chapter presents the results from the interviews with the four researchers. Based on the opinions, needs, and suggestions expressed during the interviews, a prototype was made. Such a prototype can be limitless, but this chapter covers the core requirements to a web-based collaborative system. All interviews have been conducted in Norwegian, and quotations from the participants are translated in both this and the next chapter. The presented prototype is further presented in 4.2.

4.1 Interviews

4.1.1 Data Collection

The data used in medical research is often gathered from the patients' medical record. Some of the researchers study mainly their own patients, and hence find the patients by searching through the records. Of course, the patients have to agree that their record is used for medical research, and other permissions must be in place before actually using it In the record, several different data values are available. Personal data contains info about the patient such as name, gender, and date of birth. In each record note, the physician has written free text, which may contain info about diagnoses, tests, prescriptions and a subjective evaluation of the patient during their visit. Test results are also a part of the record, like the result of a blood test.

In addition to the mixture of subjective and objective values in the patients' record, the researcher can use several other data collection methods. Questionnaires are mentioned by several of the researchers, as it often results in quantitative data that can be scanned and analyzed relatively easy. The more time the researchers have to spend on manual plotting of data, the less time is used on the actual research, so it is of great interest to researchers to automate the data extraction from records as much as possible. However, researchers also want to collect data that is not in the records, and there often is a need to interview patients, relatives or other health personnel related to a patient, or there can be a need to do some other tests or other activities.

In addition to the data sources mentioned above, we have also discussed the possibility of using a quite different, but potentially valuable source; blogs. Patients, or in this case, relatives, write extensive reports about the progression of the disease over time. In addition to the data one can find in the record such as diagnosis and medication, the blogs reports about the more subjective experience, both from the patients and the relatives point of view. How do they tackle the situation? How is their daily life affected? Do relatives become so affected that they themselves get ill? Several questions of that type can be answered using information from the blogs.

The researchers responded quite differently when discussing the possibility of using text from blogs. On one side, it was claimed that it was definitely interesting to combine several methods and types of data, such as blogs. Gaining other perspectives of the disease, such as the relative's subjective view of it, is seen as valuable, and is also a subject of some studies. However, much of the data that is made available through a blog is also available from other, more reliable sources. For instance, info about diagnoses is also a part of the health record; hence there is no need to view a blog to find that information.

A drawback with the blogs is that it is hard to assure the quality of data. Relatives could just as well write a wrong diagnosis or name of a medication. One of the researchers also mentioned that how representative the blogs are is hard to say. Do they represent the worst-case scenarios, or is it closer to a cross-section of the population? The answer is probably that the data you get from blogs is skewed. However, one of the researchers mentions that the blogs can have value "*as they are*" - if the researchers is aware that the data might be skewed, incorrect and incomplete, they still can find valuable information there. For example, a nurse could use the data in a qualitative social welfare study. Even though there are several arguments of why not to use blogs as a data source, they do have a potential as showed in the example of PatientsLikeMe, where a new treatment to ALS was suggested through the use of an illness network. Consideration must be taken as for privacy and quality of data, but looking aside of that, the potential should be further explored.

4.1.2 Register

Several of the researchers expressed the need for a dementia register. One of them states that a prospective register started in Nord Trøndelag in 2009, and will run until 2028. Work around the register concentrates on collecting data, and assures the quality of the collected data. The quality assurance is time consuming work, as 2 specialists have to work together, manually going through records to assure, for instance, that the correct diagnosis is set. The goal of the register is to combine the register data with data from another register, the HUNT register. This will result in large amounts of highly valuable data. HUNT[3] is a health register in Nord Trøndelag which has collected data from over 100 000 patients in three different surveys. As stated by one of the researchers, you can never get enough data, because many patients are irrelevant after applying inclusion-and exclusion criteria to the selection of patients. A dementia register connected with data from HUNT will therefore provide an excellent opportunity to do research on larger populations than would otherwise be available.

The drawback using such registers is that the available data values might come in short, depending on the topic of the research. It was stated that collaboration with other researchers was difficult when studying specific topics. Researchers focus on different fields, hence the actual data values gathered from each patient can be insufficient in relation to a specific field of interest of another researcher. In the register however, a core set of data for every patient is provided. In addition more specific fields are available for some patients, and unavailable for others. Even though this leads to a great challenge, the possible benefit of having such registers available is great for several researchers.

---

[3] http://www.ntnu.edu/research/research_excellence/hunt

As mentioned by one of the researchers, a crucial success criterion for a register is that there is no extra work for the practitioner when collecting data. A more automated way of collecting the data hence seems to be the way to proceed. Imagine all values from every test ever taken on a patient automatically getting fed into the register - this could be valuable for a researcher, for example when searching for biomarkers, which was an important task for one of the researchers. A relationship between a rise in a specific value (for instance from a blood test) over time and the presence of the dementia diagnosis years later, would be possible to discover using such a system. Obtaining the earlier diagnoses was mentioned as an important task within the field of dementia research and treatment. Even though not all researchers would have the desired data fields available, the potential benefit for the other researchers is great, so it is an option that should be further developed.

Even though the HUNT register contains large amounts of data, there is always a potential to combine it with several other data sources. Data from other health surveys such as Tromsøundersøkelsen and HUBRO are gathered and combined in the CONOR project (Magnus et al., 2003) , which contains health data and blood samples from about 185 000 patients. This represents a huge contribution to medical research, and especially causal research. Other existing registries such as the birth, cancer, and death register could also provide a valuable addition to a dementia register.

One of the researchers pointed out one drawback with pooling files from different sources together. Because of the continuous quality assurance at HUNT's register, the data can change. Because of this, the data in such pooled files gets outdated, and no longer represent HUNT's official data. The researcher hence suggests that each register keeps the data at site, and leaves the pooling to the researchers who uses the data, and who needs to pool data from several sources. Another issue is the problems that arise while trying to combine data from several sources when the data is almost the same, only separated by small varieties in the variables. Even within Norway, this structuring of data is a problem because of lack of standardization. Such problems must be overcome if the project is to be fully successful. The use of an onthology which structures such data could be a solution.

4.1.3 Process

Hypothesis

According to the researchers, the idea for a hypothesis can come from several different sources. It might be from literature, theories, everyday clinical work, or discussion with other colleagues. Sometimes the possessor of an idea doesn't even have time to study it, and has to pass the hypotheses on to a colleague. Therefore, the tool has to support sharing of hypotheses among researchers, and collaboration regarding them as well.

Several of the researchers mention that the hypothesis is always ready before they start studying the data material. The reason for this could be related to issues such as gaining access to the data. The owners of registers require a thorough application with a well-formulated hypothesis in order to get access to data. Access is then only given to the requested variables, as opposed to the entire data set. One of the reasons for this is to assure privacy and anonymization. However, one could imagine researchers getting access to a number of variables - sufficient to do their tasks and to discover new relations, but still limited so that the patients remain unidentifiable. This way, the researchers would have been able to browse through the data and gain another perspective of the data, possibly leading to new hypotheses and discoveries. Several tools that support such browsing already exist, and have been presented earlier in this report.

One of the researchers will only recruit patients from her own clinical practice, and does not include patients from other researchers. The reason for this was explained to be that when researching for a quite narrow area, the data collection from other researchers is not thorough enough for that particular field of interest. Hence, there is a need for the researcher to be able to collect and import their own data, and not just rely on data collected within registers. The tool is supposed to support qualitative research as well, which is another implication why several different data sources must be importable into the system.

Overview of the Data Material

Even though the researchers don't browse the data material as expected given the work of Shneiderman (2002) and Spence (2007) that was presented earlier, they do have to gain an overview of the available data. This is an important part of the research process, but as some of the researchers mention, tools that support visualisation and browsing don't exist to such a high degree. One of the researchers sees potential there. Today this is done, for instance, by studying the questionnaires (or collection method), including the target group for the collection. Missing values and categorization within variables are also examined. After this initial overview, which mainly is an overview over the available variables, the data is fed into SPSS or other such statistical analyzing tools.

Several tools and methods to gain overviews over the available data material are already in use. Scatterplots and bar charts are some of the possibilities mentioned. When researchers were presented LifeFlow, several comments were given that this was a tool that could do some of the job. Showing distribution of several values in sub groups is a useful way of comparing two or more groups. It can also be used to show what type of patients the data material contains - male/females, smokers/non-smokers, age, or whatever variable is considered interesting. However, there are a tremendous amount of variables to choose from, and all of them can certainly not be presented in this way. To keep the presentation clear, a set of desired variables must be chosen, and also the way they are arranged. NESSTAR provides such a solution, where all the available variables are presented in a "shopping chart", where the researchers can go through the available variables and choose the ones of interest.

The researchers should be provided with a tool that provides an overview also of the actual data, not just the available variables. Several methods to do this are provided in earlier chapters (e.g. KNAVE, Vistitors - see chapter 2). Choosing which patients to include is often done before feeding it into SPSS. Wrong selection of patients is, according to one of the researchers, a common source for faulty results. The choice of patients should therefore come in early, and key information about the chosen patients should be presented. This could be a way of eliminating outliers and faulty values at an early stage, minimizing the change of wrong results, or extra work when removing them later in the process.

Once the data has been fed into SPSS, the analysis is gradually made more advanced, and the dataset is adjusted for outliers etc. One of the researchers really sees the potential for improvement during this phase, as visual presentations are very useful to examine datasets prior to the analysis.

When conducting quantitative research, one usually compares groups of patients, not individuals. The level of zooming needed to obtain an overview and analyze data, depends on the hypothesis. Some requires a population-wide picture, while some requires looking at narrower groups of patients, often in relation to each other. However, some of the researchers reported that they sometimes had to zoom in on outliers and other suspicious values, in case they needed to let the patient out of the data material. This means that a visualization tool must support several levels of zooming, from showing a population, to showing a single patient or data value.

Some of the researchers reported a need to see the time-perspective, where one can see the development of different values over time. Especially researching patient trajectories, causal research, and research regarding symptoms was mentioned as fields where the time perspective could have an important role. When researching for the cause of disease, it is important to be able to go back in time and see which patients in the population actually got sick, and who remained healthy. Using data from a register can accomplish this task.

Cooperation During Analysis

Some of the researchers explained that cooperation during the analysis phase was done through actually sitting in front of the same computer. This seems quite unnecessary when thinking about the rapid development of technology today. A system that is supposed to support the researchers should definitely support cooperation during the analysis phase. One way to support this is by allowing temporary results to be passed around along with the data set, and possibly some comments. Also article manuscripts are sent around on e-mail to get feedback. This could possibly lead to several versions of the same article, each with small changes, leading to a bothersome fusion process for the responsible person. Thorough file management is necessary to avoid such possibly erroneous work. The possibility to share files like temporary research results and article proposals could also be relevant for other documents. Sharing of knowledge and experience was mentioned as important activities within a research group as well. This capturing of knowledge should be an integrated part of a tool to support a researcher’s work.

Statistical Analysis

During the actual analysis, a statistical analysis tool is used for quantitative research and several options were mentioned by the researchers, such as: SPSS, Excel, STATA, and EPI Info. SPSS was mentioned by all of the participants, and seem to be a widespread and acknowledged tool. One of their arguments for using SPSS and other statistical tools rather than a visualization tool during the analysis, is that the results have to come from statistical methods for a quantitative researcher to end up with valid results. During the analysis, competence, knowledge, and previous experience of the researchers are the most important tool. Some stated that the most important clinical analysis is done before the data is fed into any statistical tool.

The researchers stresses the importance of using statistical measures during the analysis - p-value, numbers and other objective values are key components for a valid research result, while graphs and other visual presentations are more a presentation tool after termination of the analysis. Graphs and visualizations are great to prove a point, but for the actual analysis tables and statistical measures are crucial. This was a strong declaration of why SPSS should be kept as a tool, and instead support tools should provide more visualization tools earlier in the process as well as in the finishing phases of the research.

One of the researchers, however, mentions the need for support when choosing correct statistical methods. As stated, "*We are doctors, not statisticians*". The doctors should be allowed to be doctors, and not having to work as an educated statistician as well. Some knowledge within statistics is of course necessary when doing research, but support might some times be needed when choosing the correct method. There was mentioned that a lack of statistical knowledge previously had led to an opposite result than expected. In this particular case, the researcher was aware that the outcome was wrong, but some tools to avoid such mistakes in the future should be provided. A greater degree of visualization of the analysed data and the methods that are used on it hence seems to be needed. Giving the researchers a visual overview of the analyzed data, and the methods that has been used on them could prove to be of use.

4.2 Prototype

The prototype of the medical research tool is based on knowledge gathered through the interviews, in addition to existing theory and experience from other related projects. The intention of the prototype is not to give the impression of a complete research tool, but rather serve as a point of discussion to gather more information about how such a tool should evolve. . The prototype has been kept simple as a paper-based mock-up so that the researcher more easily can get an overview over the concept, focusing on main concepts. In further iterations, the interaction with a working prototype would be the next logical step.

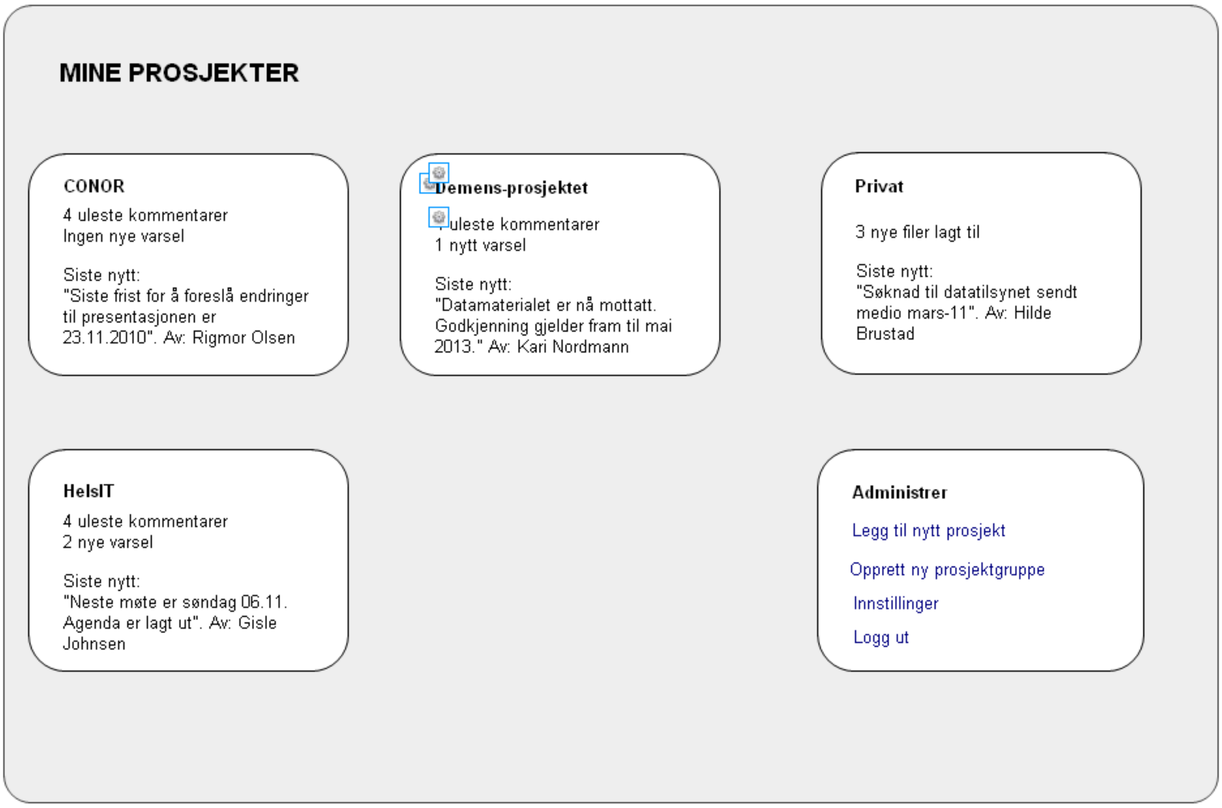
During this first round of interviews, a part including participatory design was initially planned but later dropped because it was too difficult for the participants who had trouble imagining functionality. Because of this, it was necessary to wait until a prototype was developed based on the results from this first round of interviews. This way, a concrete suggestion and "something to point at" could be the main topic of discussion during the second round of interviews.

The web-based collaborative system proposed focused on giving an overview of data sets, and eased the collaboration between colleagues through a shared workspace. Awareness regarding the status of the project and what other colleagues are doing is an important part of the prototype as it supports the collaboration process. The actual data collection is left out of the research tool. For this prototype, it has been assumed that data is already gathered and stored in files. The tool is supposed to support the collaborative, coordinating activities of a research group, as well as the examination of data to get an overview of what is available. If actual statistical analysis is needed, it has to be done in SPSS or other suitable tools, as this prototype is supposed to support both quantitative and qualitative research.

The following are an overview of some of the main functionality provided in the system, presented through the prototype. During the presentation for the researcher, some additional screens were shown in order to give an impression of the interaction with the system. A detailed picture of each of the presented screenshots can be seen in Appendix A.

Screen 1: Overview over the researchers’ projects

One screen must provide a short overview of all the projects the research is involved with. Getting to know the status of each project is important, so each project contains information about new messages, comments, or files in addition to the latest news for the project. A screenshot is provided in Figure 11.



MINE PROSJEKTER

CONOR
4 uleste kommentarer
Ingen nye varsel

Siste nytt:
"Siste frist for å foreslå endringer til presentasjonen er 23.11.2010". Av: Rigmor Olsen

Demens-prosjektet

uleste kommentarer
1 nytt varsel

Siste nytt:
"Datamaterialet er nå mottatt. Godkjenning gjelder fram til mai 2013." Av: Kari Nordmann

Privat

3 nye filer lagt til

Siste nytt:
"Søknad til datatilsynet sendt medio mars-11". Av: Hilde Brustad

HelsIT
4 uleste kommentarer
2 nye varsel

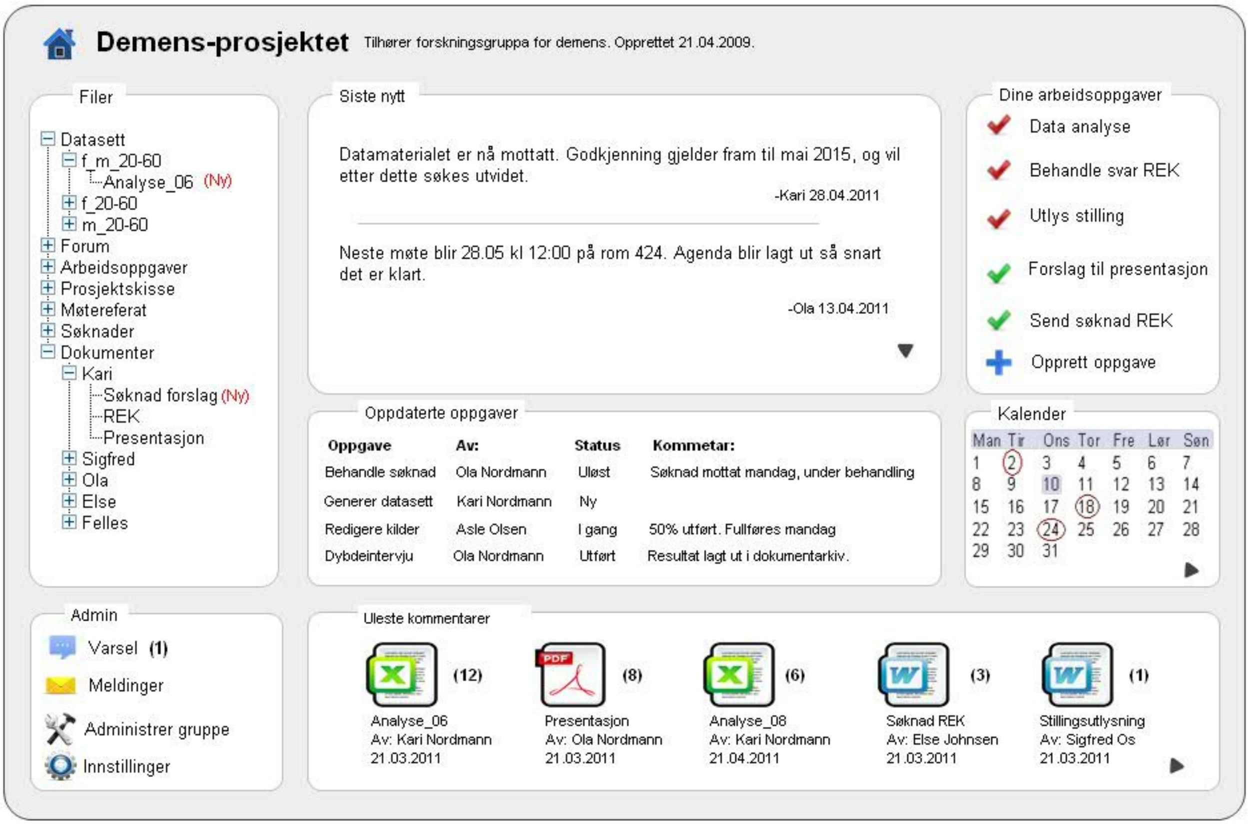
Siste nytt:
"Neste møte er søndag 06.11. Agenda er lagt ut". Av: Gisle Johnsen

Administrer

Figure 11: Projects of which a researcher is involved in.

Screen 2: Overview over a chosen project

Each individual project has its own page which gives more thorough information (see figure 12). Latest news, status updates, and access to the shared files are given. The possibility to share files through a shared workspace is one of the main functions of the system. Communication around the shared documents is visualized in this screen, by making the user aware of the latest actions related to certain elements. At the bottom, documents with unread comments are shown. As one can see, awareness is also supported by presenting the status of tasks executed by other researchers. A calendar for the project with important meetings and other events are also provided, in addition to an overview of the researchers own tasks.



Datamaterialet er nå mottatt. Godkjenning gjelder fram til mai 2015, og vil etter dette søkes utvidet.

-Kari 28.04.2011

Neste møte blir 28.05 kl 12:00 på rom 424. Agenda blir lagt ut så snart det er klart.

-Ola 13.04.2011

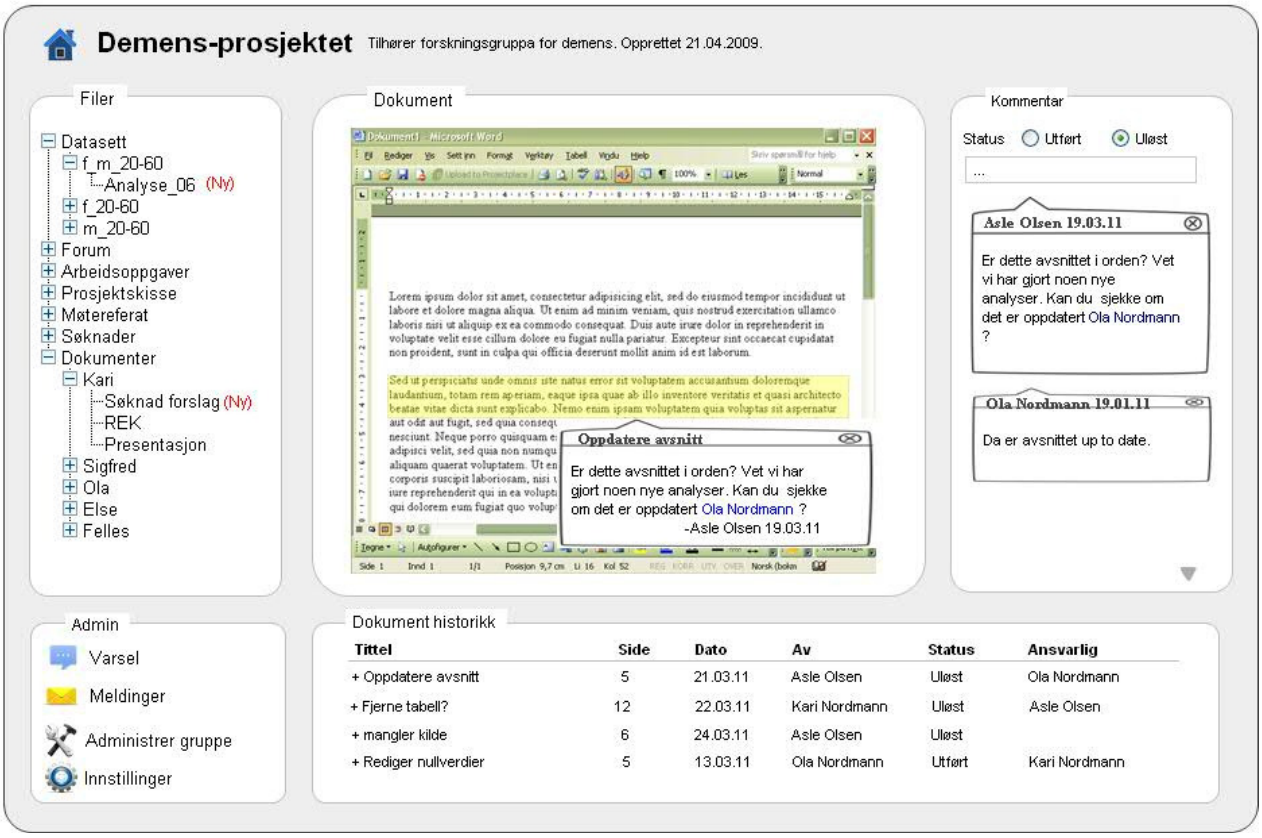
Oppgave	Av:	Status	Kommetar:
Behandle søknad	Ola Nordmann	Uløst	Søknad mottat mandag, under behandling
Generer datasett	Kari Nordmann	Ny	
Redigere kilder	Asle Olsen	I gang	50% utført. Fullføres mandag
Dybdeintervju	Ola Nordmann	Utført	Resultat lagt ut i dokumentarkiv.

Man	Tir	Ons	Tor	Fre	Lør	Søn
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

Figure 12: Overview of a specific project.

Screen 3: Notification and tag function

From the last screen, one can see that the researcher is made aware that a notification ("varsel" in Norwegian) has been given. The notification implies that a colleague has tagged the researcher - in this specific case, in a comment for a document. The tag and notification function can be used to make colleagues aware of a comment, task, or file that might be of interest - in this case, a question about a paragraph in a document has been asked, as seen in figure 13. Often, researchers involved in a project work from various locations and with different time schedules. Functionality that supports asynchronous communication is a necessary part of a research tool that is supposed to enhance cooperation. The document sharing and having the possibility to discuss written texts (e.g. comments, ideas, etc) , is also included in the prototype in order to support the way the researchers actually work. Such possibilities reduce the need to meet the colleagues face-to-face, or the alternative of having such discussions in everlasting e-mails with referrals to paragraphs in different documents. However, occasional face-to-face meetings are often necessary in addition to online communication, in order to have effective communication.



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia conseq

Oppdatere avsnitt

Er dette avsnittet i orden? Vet vi har gjort noen nye analyser. Kan du sjekke om det er oppdatert Ola Nordmann ?
-Asle Olsen 19.03.11

Status ○ Utført ● Uløst

...

Asle Olsen 19.03.11

Er dette avsnittet i orden? Vet vi har gjort noen nye analyser. Kan du sjekke om det er oppdatert Ola Nordmann ?

Ola Nordmann 19.01.11

Da er avsnittet up to date.

Dokument historikk

Tittel	Side	Dato	Av	Status	Ansvarlig
+ Oppdatere avsnitt	5	21.03.11	Asle Olsen	Uløst	Ola Nordmann
+ Fjerne tabell?	12	22.03.11	Kari Nordmann	Uløst	Asle Olsen
+ mangler kilde	6	24.03.11	Asle Olsen	Uløst	
+ Rediger nullverdier	5	13.03.11	Ola Nordmann	Utført	Kari Nordmann

Figure 13: Shared document with comments on a specific paragraph.

Screen 4: Overview of dataset

The prototype contains functionality to support more than the communication and cooperation aspect.

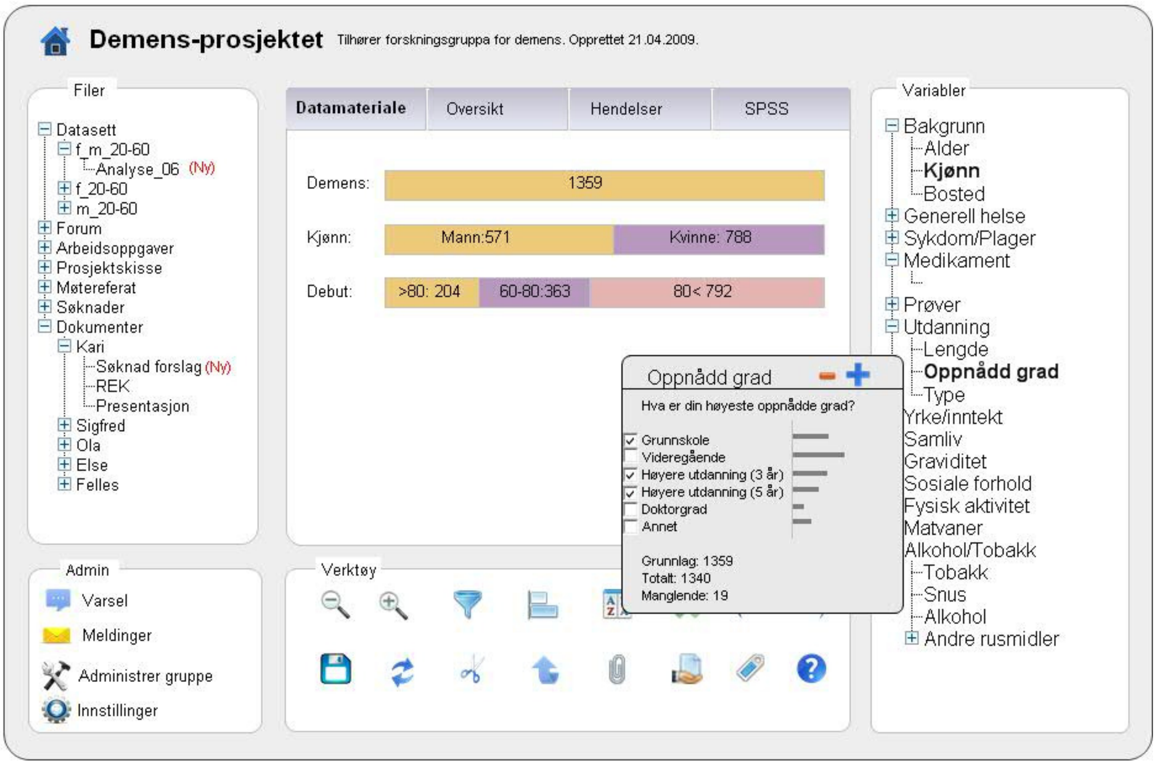
The screen where the researchers examine and analyze the data is divided into four tabs:

1) Get an overview of the available data material
2) Examine the current patients of interest
3) Look at event sequences related to the current patients of interest
4) Do statistical analysis using SPSS or other tools

Tab four is not provided in this overview, as the statistical analysis of the data is not done directly through the system, but rather through familiar, already used tools such as SPSS. A seamless integration with such tools is important for the researcher, as these are tools that are already in use and work well for them. The system is intended to be a supplement, not a replacement to such tools. Including SPSS here is an attempt to get a discussion going with the researchers as for how this integration could work.

The first tab supports the researcher in gaining an overview of the available dataset. The researcher needs a way to get an overview of what's available, and what's not in order to judge if the data material is adequate. A natural starting point, as mentioned in the result from the interviews, is to give an overview of the dataset and the patients' key variables, such as age and gender. It is also possible to see what variables the dataset contains, and important metadata such as collection method and number of missing values.

The screen provides a graphical visualization of the dataset (figure 14). Several different views should be provided in order to fit in with each individual researcher and the current data of interest. On the bottom, a toolbox is available with several useful tools, such as sorting and zooming. Throughout the process of handling the data material, the toolbox is available for the researcher to use to manipulate the data.



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

Demens: 1359

Kjønn: Mann:571 Kvinne: 788

Debut: >80: 204 60-80:363 80< 792

Oppnådd grad

Hva er din høyeste oppnådde grad?

☑ Grunnskole
☐ Videregående
☑ Høyere utdanning (3 år)
☑ Høyere utdanning (5 år)
☐ Doktorgrad
☐ Annet

Grunnlag: 1359
Totalt: 1340
Manglende: 19

Figure 14: Gain overview of a dataset

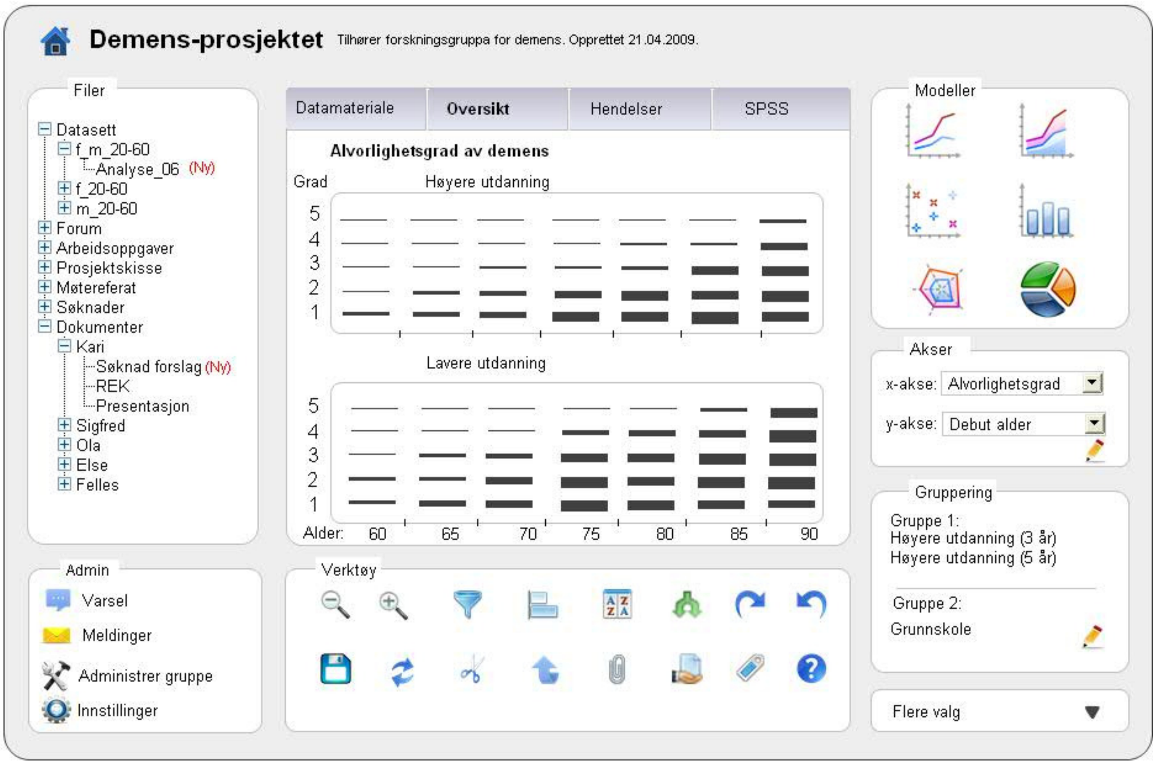
Screen 5: Get an overview over chosen patients

Before starting the statistical analysis, a researcher needs to do some further investigation of the data. The data is presented graphically, so that the researcher is able to examine the data in an effective manner. There are several possible tasks during the data analysis, some of which is presented in the list below.

- filter
- sort
- align
- rank
- zoom
- tag
- comment on a data set
- compare patients or group of patients
- change view (type of diagram, change denomination, etc.)
- find more similar patients

- store and share a search, data set, or graphical visualization
- undo, redo
- show statistical measures

The tool should support several different tasks and visualization methods, to fit in with each of the researchers' workflow. Some are doing qualitative research, whereas other is doing quantitative research. Also several different kinds of data are used, so a tool must support a variety of tasks and visualization techniques. Presented here is a quantitative research task - it could just as well be showing a single patient through an interface that is something similar with Lifelines or PatientsLikeMe that was presented in chapter 2. A screenshot for examining current patients of interest is provided in Figure 15.



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

x-akse: Alvorlighetsgrad

y-akse: Debut alder

Gruppe 1:
Høyere utdanning (3 år)
Høyere utdanning (5 år)

Gruppe 2:
Grunnskole

Figure 15: Examining current patients of interest

Screen 6: View event sequences

The third tab provides an interface that supports features regarding event sequences. As shown in Figure 16, a researcher can see a series of events for one or more groups of patients. In this specific case, the prototype visualizes the event sequences for two distinct groups of patients, one with higher education, and the other without. The two graphs show that there might be a correlation between level of education and progression of the dementia disease, which could be further investigated with statistical analysis. This way of presenting the data visually, allowing the researcher to "browse" through the data, might lead to the discovery of new hypothesis as the comprehension of the data increases.

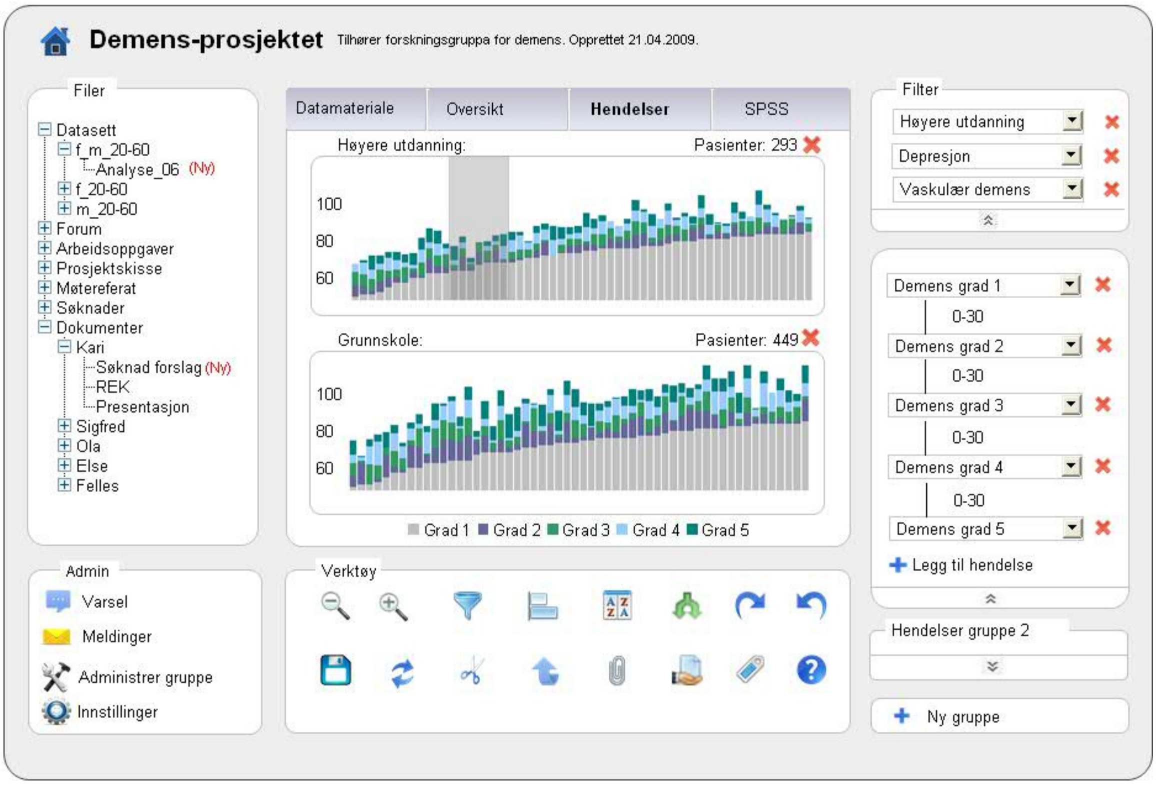
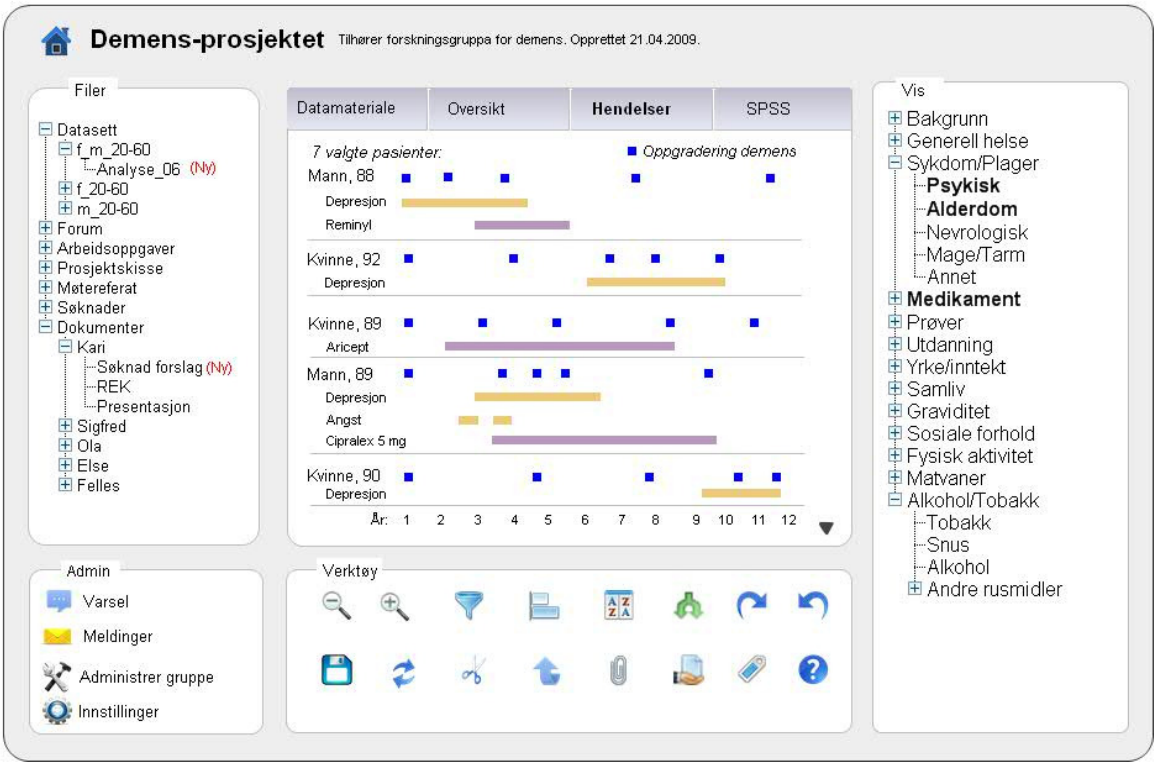


Figure 16: Viewing event sequences.

Screen 7: Closer look at chosen patients

If fig 16, some patients that the researcher wishes to take a closer look at, is marked. In fig 17, the system has zoomed in on these patients, providing a more detailed view of the patients. The research can specify the level of detail. There is also the possibility to zoom even further in, and take a look at a single patient, using the same picture. Here, several events regarding the 7 patients are shown - duration of events and events happening at the same time might reveal something about the data that was previously unknown.



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

Figure 17: Taking a closer look on specific patients.

5. RESULTS INTERVIEWS ROUND 2

During the second round of interviews, the prototype was presented to two of the participants, followed by a discussion about the concept, specific functions, and possible extensions and changes. The reactions and suggestions expressed during the second-round interviews are presented in this chapter.

5.1 Cooperation and coordination

The concept of cooperation support and coordination was met with positive response. The "latest news" section was considered useful, as it can contain status updates, information on courses that have been participated in, or as in this case, that the dataset is added. *"I believe that such a web-based cooperative system could work, definitely",* said one of them, whereas the other said *"when communicating, to be able to get a facebook-like update on what is going on in different projects, I think that is smart".*

For the first version of the prototype, a calendar function was suggested in order to mark special events and meetings. One of the researchers questioned the way the calendar would integrate with other calendars, as many of them often used one calendar to coordinate all their activities. The other expressed that a tool that would help organizing all the activities going on, would be very useful. Also the availability for the system on other platforms such as iPhone was asked. Getting an overview and having control on all projects, messages, tasks and emails at once was a difficult and time consuming task. Especially e-mails was mentioned as time consuming and hard to keep an overview of, since they recieved huge amounts of e-mails every day.

A calendar function in it self seems to be useful to both participants, but the integration must be further examined. There should be possibilities of integration that do not require the researchers to change the way they work today - unless they want to. Forcing users to use the system in a certain way or instead of other systems will only lead to less approval and resistance to adoption. However, sorting emails and doing a part of the screening for the researcher, seems to be desired as the number of e-mails received is great.

There is often cooperation around different documents, and the coordination of this is not easy. Today, this is done in two ways. Either by sending to all co-authors at the same time, collecting feedback, and integrate all the new versions of the document into one. The other option is sending the document to one co-author at a time, and wait for response before it is retransmitted - this is of course very time consuming. As one of the researchers says, "*there is a potential of improvement in efficiency during such rounds*". Getting to know what is going on with a document, without having to send several emails would help, especially if there is any possibility of showing deadlines as well.

The solution with comments and tagging in documents was seen as progress during the work that requires cooperation. Additional functionality was also suggested, that didn't come up during the last series of interviews. A Skype/go-to-meeting function was suggested, where a group of researchers could all see a document, and at the same time speak via a calling function, similar to Skype. This way, one responsible person could go through a document while talking to the colleagues (who could also view the document), marking individual sections with a comment, and possibly "tag" the responsible colleague in the section of current interest. This way, there is no doubt about who is responsible for what parts, the distribution of work is made more efficient, and discussions around a document are taken with all those involved present. It also allows for easier integration of colleagues that works from other locations, which is a common work situation in today’s world.

Even though great potential was seen in the solution around documents, a challenge was pointed out. *"It has to be a good solution for it to be adopted - everyone has to agree to use it"*. This is not only a challenge for the document handling in this system, but rather in the whole domain. As expressed, hard prioritizing has to be done as for which technologies to adopt while working within the domain.

The issue of transparency came up during the interviews, as one of them mentioned the possibility to view what colleagues are currently doing. In the prototype, this is represented through a list of currently updated tasks from other researchers. However, a more extensive view on it was considered helpful. The possibility of allowing each participant in a project create their own profile was also mentioned during another discussion. These two wishes could successfully be combined with the "go-to-meeting" function mentioned above. Each participant in the project could be presented on the project overview, so the others could see who is currently online. Each participant could then have its own profile page, where information about themselves (name, contact info and other necessary information for a manuscript to an article), a list of tasks along with the status of those tasks (similar as shown in the first prototype), and a list of their publications. Displaying "on-line status" when others are available, leads to another well-known feature; instant messaging. One of the researchers mentions that it could be an effective way of getting to know what other people are doing, for instance wondering what they are doing with a specific document since it is locked at the time, to avoid unnecessary duplication of work, or to get a quick clarification on what to do with a specific patient.

Sharing knowledge is another task that a CSCW system should support. Each researchers profile page could provide a part of this experience and knowledge sharing. Several options exist, but a way of presenting the experiences of each individual could be through a personal blog. Of course, using such a blog is optional, and no extra workloads should be put on the researchers. Only functions that are considered useful are adopted, so hopefully such a function could provide a way of sharing the information. Other options include the use of forums or wikis, in addition to using the already suggested commenting, - and notification system. By writing blogs or forum posts, others are able to comment on them, and maybe a "recommend" function could work.

Not only knowledge and regular documents are shared among the participants in a project. One of the researchers mentioned that they had a common literature base, where links to relevant articles could be found. Such a litterature base should be included in the prototype. Once the literature is a part of the common workplace as well, one should of course be able to search it, but also other options such as recommendations by colleagues and related or linked documents could be an option to further discuss.

5.2 Data overview

In addition to the collaborative activities, a part of the prototype displays an overview over the available data using different visualization techniques. When discussing the possibility, the first impression seemed to be a bit mixed. On the one hand, one of the researchers could not imagine where the data in the figures presented would come from. A register does not provide data in such a format, as such data cannot be open and accessible to everyone. In addition, researchers who want access to the material have to apply, and are granted access only to limited parts of the data material, according to the application. This limitation of accessibility is both because of privacy issues, and to protect other researchers. If for example a PhD student wants access to a part of the data, then no other researcher with a similar hypothesis or research area is allowed access. In these cases, one would prefer if several researchers cooperated instead of competing with each other.

Even though the first impression was somewhat mixed, after discussing the possible use of such a visualization, the researcher seemed more positive, and explained that it can be useful to see what data is available. Once they get a hold of the data, the work of acquiring an overview starts. There are tools available for this in statistical programs, but one problem is that they are often poorly visualized. Having a good overview before starting more advanced analysis can be useful, and perhaps prevent mistakes such as selecting the wrong variables so that the result from an analysis comes out the opposite than what they should.

> *Often, you start with a table, gender, age, disease/not disease, make figures, etc. That is what you do before starting regression analysis. This system can replace or simplify a part of the first overviews. Addition or replacement to the first analysis of the material... Interesting.*

The other researcher was more sceptical to developing yet another statistical tool. The proposed system has another point of view than the other, more traditional statistical tools, in that it focuses more on the overview before the actual statistical analysis. Even so, the threshold to learn how to use new systems is high, because of already existing time pressure.

The added value hence must be greater than the extra time and effort in learning more systems.

The connection with SPSS or other statistical tools were also discussed. As the researchers seem satisfied with these tools, there is no need to change them just in order to make one mega-system. A way to allow exchange with these tools should be provided, so researchers can keep using familiar products. Avoiding disturbing the researchers established workflow too much is crucial to gain their acceptance. A tool that requires too much change too fast has higher probability of not being adopted.

Another issue discussed, was that researchers handle research files differently. For instance, in the case of making selections from an available dataset, some researchers store the new dataset in an individual file, whereas others only use filtering on the main dataset. People have different ways of working, and a tool must support this. One suggested solution was to keep the research tool as a file archive and cooperation-support tool instead of including the visualization part. This way, they don't necessarily have to be forced into one way of working, but can rather download the necessary files. *"Having a base where all the necessary files are is a captivating thought"*. However, security and availability must be considered. Even though the data is deidentified, it can not be shared to just anybody.

An important point that was discussed is the issue of system adoption. Getting systems adopted within the healthcare industry is a difficult task, as there is hard prioritizing of what they are able to adopt. Some tools are already available for making overviews over variables, creating tables etc, but for some reason it is not prioritized. The researcher claimed that it was not because of low usability, but there was still a threshold to start using such tools. This is as already explained probably because of a lack in time, and that they already have tools that do the work "*good enough*". Developing for adoption is therefore a key issue for the success of a research tool. The perceived benefit must seem larger than the effort of learning and adopting new systems.

6. DISCUSSION AND ANALYSIS

The results from the interviews and the presentation of the prototype to researchers has resulted in a set of core requirements for a dementia research support system. The requirements changed drastically during the two rounds of interviews, and presented here are the requirements that should be the basis for future work. This chapter presents the discussion and analysis related to the research questions. First, the current situation is described, followed by a discussion of how dementia research work can be supported and made more efficient using a web-based collaborative tool. Finally, the different potential data sources are discussed.

6.1 The research process today

Researchers within dementia have, as do other researchers, very busy days with several activities going on at the same time. A time-consuming task was found to be the organization of all those activities; prioritizing and planning were mentioned as key tasks. The planning is done mostly using e-mail. Several of the researchers mentioned they got a lot of e-mails, hence a way of sorting and filtering uninteresting email, or possibly lower the number of incoming messages was suggested. Communication with colleagues was also mentioned as a great part of their job. However, much time is spent getting to know what others are doing, and coordinating their work. Some way to ease and make the communication more efficient is wanted.

As for the actual research process, an important part of it includes data collection. Collecting the data, assuring the quality, and feeding it into a register are important but time consuming tasks. Especially feeding the data into a register was mentioned as a routine task that could be more effective.

Making the hypothesis is also a part of the process. The hypothesis comes from literature, theories, everyday work, or through discussion with colleagues. The hypothesis is sometimes shared, for example when one individual researcher doesn't have time to do the research himself, he might participate in the formulation of it and then pass it on. Knowledge sharing should be a part of the system, in order to support this information exchange. Sharing knowledge is one of the main features of medicine 2.0 approaches. Also other areas such as literature, publications, and presentations should be made available to the research group to support the collaboration and knowledge sharing. Eysenbach (2008) mentioned oppenness and sharing between apomediaries as two of the main concepts of medicine 2.0, meaning that researchers should be able to share, recommend, and guide each other through the vailable infomration. The results gathered shows that the researchers are open-minded and willing to participate in such work, as they consider it helpful.

After the data is collected, the process of getting an overview starts. This can be done in several ways, for instance in studying the collection method, available variables, or the participants themselves. The statistical programs offer possibilities to get a visual overview of the available data, but these are not so commonly used during the examination phase according to some of the researchers. Making graphs and other visualizations takes time, and are for some reason not prioritized. Instead, tables that shows the number of, for instance, male/female and smoker/non-smoker are used.

For the statistical analysis, standard products exist and are used. SPSS, STATA and Excel were mentioned as most frequently used. During the analysis, a lack of competence in statistical methods was mentioned as a problem under some circumstances. Sometimes, mistakes are made and a variable could be selected that was unintended and leads to wrong results. Visualization of the data might prevent such mistakes, as well as providing extensive support in the earlier phases of examining the data (Wang et al., 2010) and proposing hypotheses. Kuhlthau mentions uncertainty as increasing during these first phases of research (1999) Visualizing the data might reduce this uncertainty. Even though visualization helps for some problems, and possibly makes the exploration phase more intuitive, it has some drawbacks. The main problem is the problem of adoption. There are already so many competing tools available, so why learn yet another one? The perceived benefits must be higher than the energy to learn and use yet a tool. The question arises as if such a tool can provide enough benefits for it to be adopted

If such a system is to be made, it must either replace other statistical tools by a better alternative, or provide a close interaction with them, for example by exchanging files between them. The integration must support several tools, as researchers often have a favourite, and prefers to continue with a tool they know well.

The result of the research is often an article or report to a professional journal or professional conference. For the professional journal article, the research group often creates a graphic visualization in order to show the results in an intuitive way. The visualizations are then put into the article, together with tables, images and of course the text. The coordination and collaboration around documents was mentioned as time consuming, having great potential of increased efficiency. Around a document, there are often several authors, possibly with different backgrounds such as a statistician in addition to the other researchers. All the involved persons collaborate to create the document, and must be given the opportunity to propose changes to some of the elements. This process is mainly done in two ways. A coordinator can send a version to all participants at the same time, receives comments, and makes the necessary changes. The other option is to send to one participant at a time, so that the participant can make the changes directly into to article. Both those methods are unnecessary time consuming, and should be improved.

6.2 Proposed tool to support the research process

As mentioned in the last subchapter, there are several points where technology can be used to improve the research process . The system presented to the researchers mainly consisted of two parts; (1) collaboration and coordination of the research group, and (2) tools to get an overview of the available data material.

The parts of the prototypes regarding project coordination and collaboration between the researchers were perceived as very useful, with a great potential to make the research process more efficient. Especially the functions related to document handling were seen as helpful, as this is a time consuming task. A well-known web 2.0 application today is Facebook, and the possibility to tag, comment and receive notifications was attempted as a part of the prototype. The functions were very positively reviewed by the two participants, among others because they had potential to decrease the number of emails received. This is both because notifications could be used instead of emails, and in addition, no articles have to be sent around to gather comments. This saves time that could be used on other, more important tasks than reading emails.

Another effective way of decreasing the number of email received is to gather all incoming articles into one "mailbox" stationed within the tool. One of the researchers stated that several emails were just referrals to articles, and maybe just 5 out of 20 were relevant. However, the participant wanted to receive them, to be able to get an overview over what was going on in the field. Allowing the researchers to share articles, and possibly rate or recommend them, would save a lot of e-mails. However, articles from other people than within the research group are received, such as automatic messages from the library. Some way of integrating those articles into the system should be further investigated for the function to be a success.

Another part of the system is the calendar function. It would have to be integrated with other calendars to work as intended. To be able to get an overview over both personal tasks, and the tasks of others in the project were perceived as useful, as it increases the awareness and makes it easier to coordinate the work. Another great possibility with the calendar is to organize phone meetings with several participants. Imagine integrating functionality from group video calls such as Skype with the document handling, where a project group could be able to talk, view the same document, and comment on specific paragraphs in the document in real-time. This would save time and effort regarding planning, communication and collaboration. Tasks could be assigned on the fly, discussions around a paragraph could be taken at once, and misunderstandings could be avoided. According to the participants in the study, this is a functionality that should be further developed.

The awareness of what other group members are doing was also mentioned as a helpful feature, which is mentioned as one of the key factors in CSCW systems by Pratt (2004). In the prototype, this was done by having a list of recently created or executed tasks. The researchers asked to extend this task-list functionality. A way to do this is to let participants in the research group have their own personal page, where latest news related to the specific group member is kept. An overview of the members tasks, short status messages, published articles, and so on could be gathered in such a page in addition to personal information. Other opportunities also exist, for example when sharing knowledge. Writing blogs or posting presentations could be a way of sharing knowledge. However, such functionality might not be used, because of the extra workload required to actually write a blog post. It should therefore be further investigated if there is any interest in such functionality before it is actually implemented. However, it does provide easy knowledge sharing between apomediaries.

The other main part of the proposed system was to provide a graphical visualisation of the available data. The purpose of this was to make any correlations between variables more evident. Examining the data with available tools such as filtering, sorting, ranking and zooming could provide a more intuitive overview than when setting up tables, as suggested by Wang et al. (2008). Even though there is a potential for improving the task of getting an overview of the data, the functionality was received with mixed opinions among the participants. It was perceived by one of them as having potential to be an addition or replacement to the first overviews one has to get, assuming that integration with SPSS, STATA and other tools are possible. Some mistakes regarding mistakenly swapped variables and missed correlations might be avoided, and in addition, one has to know what data is in the material before grouping and analysing more thoroughly.

On the other hand, as there already exist several systems that can be used to explore the data, the need for yet another tool can be discussed. One of them mentioned that some programs already provided such overview-solutions, but that they were rarely used for various reasons. The threshold of learning new tools was mentioned as one reason, even though such solutions provide additional functionality. However, the potential from a visualization system like the one presented, are not justified by a relatively static prototype as the one shown to the researchers. The interactivity and dynamic visualization is hard to imagine, and should probably be experienced first hand to explore the strengths and benefits from such a system. This is a well-known disadvantage with low-fi prototyping, and should be taken into consideration when assessing the potentiality of such a system. Other research projects conclude that there is a need for more visualization of the data, as is already discussed in chapter 4. This is why the feature should be further evaluated before completely discarding it.

6.3 Potential data sources

The proposed prototype contains functionality to cover several of the needs expressed earlier. However, the suggested prototype only covers the part of the process after the data is collected. Other requirements will have to be included if the data collection is to be a part of the system. Several potential data sources are available, some of them less traditional than others.

The patient's health record is one obvious source of information. Several researchers have some clinical work in addition to the research, so recruiting patients from their own clinical practice often comes naturally. However, only recruiting from their own patients might be too limited for a quantitative study, so additional material is needed. This can be gathered through cooperation and sharing with other researchers, but are there any other potential data sources available?

The next step after collaborating and sharing data with only a limited group of well-known researchers are health registers. Several health registers already exist, such as the national cancer, birth and death registers. These registers might not provide enough information for a dementia researcher, so other registers need to be considered as well. HUNT is one of the registers considered in this study, as it contains information about several other aspects of a patient, including lifestyle and other variables that could be used to research the cause of dementia or other diseases. The sharing of data as done in HUNT do emphazise sharing, but does not support the Medicine 2.0 approach to openness (Eysenbach, 2008) to that great extent since researchers who wants access has to apply for it, and in addition, the given access is quite limited .

A specific dementia register is also under construction that contains additional information more related to dementia beyond the HUNT register. A merging of those registers would provide valuable data source to many researchers. However, it would not be enough for all researchers, as some require more specific variables. One of the participants argued that this was the reason why cooperation with others was difficult - different interests' leads to different needs and requirements to the data material. The use of registers hence might not be the best option for genetic researchers, but might prove efficient for other areas, such causal research. However registers do provide a large number of patients which is hard, if not impossible, to gather by a single research group who would have to devote considerable resources to data collection.

The data sources already mentioned are still collected by health professionals. This ensures the quality of data, and provides a certain type of variable. A quite different approach, is using patient-produced sources in the data material. As the medicine 2.0 approach suggests, the patient, practitioners, and researchers should find new ways to cooperate (Eysenbach, 2008). The data collection could be a part of this cooperation, either through illness communities such as PatientsLikeMe, by writing blogs, or by patient-entered health data on for instance a smart phone application, website etc.

For this study, posts from a dementia blog written by a relative, and some forum posts were presented to the researchers. This data source suggestion produced mixed reactions. One was very sceptical to using this "*private information*" in research, even though it is freely available on the Internet. In addition, the quality of the data could not be ensured. However, some of the researchers saw a potential in such information, as they provided whole new types of variables, especially regarding the subjective feeling around such a disease. For example, when relatives and caregivers become an "*extra patient*" because of the pressures experienced is a well-known issue within dementia. Blogs often provides very personal, descriptive writings about how the disease affects the lives of those surrounding the dementia patient. Getting to know how relatives react, what is experienced as comforting, and how different situations are handled could be valuable information, and is a research area of it's own. Though not that large yet, but the potential in such blogs should not be undermined.

One of the researchers suggested that the stories could be used by a creative researcher in doing nursing research. For the time being, gathering quantitative information from blogs seems to be difficult, because of lacking natural language processing and information extraction techniques. One also has to question to what degree the blogs are representational for the entire population - maybe just the worst cases are brought up through blogs. At the time being, using the blogs as they are, conducting qualitative research on the data provided seems like the best choice. However, they could be one way of including the patients closer in the process. Closer collaboration between patient and proffesionals is one of the purposes for the medicine 2.0 approach, and increased participation from the patient is mentioned as a typical Medicine 2.0 feature(Van De Belt et al., 2010).

Several other ways of including the patients in the research process could be an option, but this is not further discussed in this thesis. The belief is however that alternative data sources could provide a valuable addition to today's way of collecting data. Including the patients as suggested by the Medicine 2.0 approache should be further investigated, as they might be a new and valuable source of information, providing data that could not be gathered otherwise. One suggestion was made for letting the relatives of a dementia patient plot in data that could be seen both by a researcher and the physician. This would ease the researchers' and practitioners' work, and enhance collaboration between some user groups that have not currently cooperated in that way - just like the Medicine 2.0 approach suggests (Eysenbach, 2008).

7 CONCLUSION

In this study, four persons related to dementia research have participated. A case study has been conducted, with the main purpose of getting to know how dementia researchers work, how a web-based collaborative system can support this work, and several more or less traditional data sources for their research have been discussed. This chapter makes a summary of the main findings presented earlier, and continues by suggesting further work and the validity of this study.

7.1 Conclusion

> **RQ: Are there specific web-based, social collaboration and data visualization methods that can improve the efficiency of dementia researcher's work?**

Researchers within the field of dementia often work close with other researchers in project groups. One researcher is often involved in several projects at once, and does much work to coordinate this. A great deal of e-mails is received on a daily basis to coordinate this work. Collaboration with colleagues is an important task, especially around handling documents, making sure all co-authors get to express their opinions. There are also several tasks directed to the actual research process, in addition to the collaboration and coordination issues. Collecting data, stating hypotheses, getting an overview on the data, and analysing the data are all main tasks for the researchers. For some of this work, well-known tools such as SPSS, STATA or Excel are used. Even though the researchers could identify some shortcomings with those systems, especially related to the phase where an overview is desired, they still found them useful enough. During busy working days, the benefit provided by a new tool must be evaluated upon the effort of learning those tools.

For this study, two main approaches to support the researchers' everyday work were suggested. A tool to support collaboration and coordination seemed to be desired. The suggested approach was perceived as timesaving, and useful especially in the cooperation around documents. Getting aware of what colleagues are doing and the status of a project were considered helpful, as well as handling tasks and getting reminders of tasks through the integrated calendar. A system that follows the design of Web 2.0/Medicine 2.0 should be further developed and evaluated, as the perceived usefulness was positive.

The other part of the suggested system was a statistical tool that supports mainly the earlier parts of the data analysis as an addition or replacement to other tools. It was met with mixed opinions. Such a solution did have potential, however, the question was whether it added enough to the existing tools to be worth learning and developing further. However, understanding the full potential of an interactive, dynamic system is hard when presented with a relatively static prototype. Because of this, the need for such a tool should be further evaluated before it is thrown away.

Several data sources for the researchers are available. However, not all of them are of interest to all kinds of research. The traditional approach of collecting data from the patients' records or via schemes/questionnaires works well for many research projects since the data are collected by health professionals and the quality of the data can be assured. The possibility of merging such collected data with data from other researchers, or potentially use a health register, creates great opportunities for quantitative research. Closer follow-up of individual patients through using their record, interviewing, or other suitable methods also creates a basis for qualitative research.

Other less traditional approaches also have been suggested, such as using patients' blogs, forums, or illness communities. This approach uses data produced by the patients themselves, and cannot be quality assured to the extent health professional collected data are. However, the potential exists for using such sources in qualitative research. They provide valuable insight into how patients and their relatives react to the illness, and how they handle life. New suggestions to treatments have already come out of such research, as shown in the PatientsLikeMe project. Illness communities also have the potential of being a part of quantitative research. However, one must take into consideration that the data might not be representative as the patients participating in such communities belong to a certain patient group. Using the Medicine 2.0 approach seems to create value, as it provides closer cooperation between patients, practitioners and researchers. One could imagine patients themselves plotting data in on a Smartphone, being immediately available to the researcher and practitioner. However, such an approach must also provide some benefit for the patients to be willing to do it.

7.2 Further Work

The concept of supporting collaboration and coordination between researchers should be further investigated and developed. As a result of limited access to the researchers, only interviews were conducted in this study. Further work should also include observation of several researchers and research groups, as getting to know more about the collaborative activities is important when developing CSCW systems. The willingness to adopt such a system also needs to be further investigated, on a larger number of participants than provided in this study.

As for the statistical tool, it should be developed into a hi-fi prototype, where the participants of a study are given greater opportunity to interact with the system. Exploring concrete data that the researcher can relate to, in an dynamic way are key issues for such an assessment to give information about whether such a system should be further developed or not.

As suggested, adopting Medicine 2.0 approaches into the dementia research process should be further investigated, as the potential for added value are great. Willingness to participate, quality of data, and patient privacy should be key issues in such a study, as well as studying ways in which the patients themselves can contribute to the research through for example systems like PatientsLikeMe.

7.3 Validity of results

The work presented has been a case study of four participants related to dementia research. As the number of participants has been low, the results might not be representative to other researchers. However, other studies support several of the findings. Observation has not been conducted due to limited access to researchers. Observation should be a part of a case study, especially when collaboration and cooperation are key issues as it has been in this study.

During the interviews, the interviewer attempted to be as non-threatening as possible. Even so, the participants may have held back information, and opinions given may have been interpreted in another way than they were meant. However, several of the participants expressed similar opinions, which support the belief that they gave their sincere meaning.

References

AVISON, D. E. & MYERS, M. D. 2002. *Qualitative Research in Information Systems: A Reader*, Sage Publications Ltd.

BAECKER, R. 1994. *Readings in GroupWare and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration*, Morgan Kaufmann Publishers Inc.

BAR, D., SHAHAR, Y., AIZENBERG, M., BOAZ, D. & TAHAN, G. Year. KNAVE II: the definition and implementation of an intelligent tool for visualization and exploration of time-oriented clinical data. *In:* AVI '04: Proceedings of the working conference on Advanced visual interfaces, 2004 Gallipoli, Italy. ACM, 171-174.

CORNFORD, T. & SMITHSON, S. 2006. *Project research in information systems: a student's guide,* London, Palgrave.

ENGEDAL, K. 2003. Demens og demenssykdommer. *In:* ENGEDAL, K. & WYLLER, T. (eds.) *Aldring og hjernesykdommer.* Akribe Forlag.

EYSENBACH, G. 2008. Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness. *Journal of Medical Internet Research* [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2626430/ [Accessed 13.05.2011].

FAILS, J., KARLSON, A., SHAHAMAT, L. & SHNEIDERMAN, B. Year. A Visual Interface for Multivariate Temporal Data: Finding Patterns of Events across Multiple Histories. *In:* IEEE Symposium On Visual Analytics Science And Technology, 2006 Baltimore, MD, USA. 167-174.

FROST, J. & MASSAGLI, M. 2008. Social Uses of Personal Health Information Within PatientsLikeMe, an Online Patient Community: What Can Happen When Patients Have Access to One Another's Data. *Journal of Medical Internet Research* [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2553248/ [Accessed 30.11.2010].

GOBLE, C. 2005. Using the Semantic Web for e-Science: Inspiration, Incubation, Irritation. *In:* GIL, Y., MOTTA, E., BENJAMINS, V. & MUSEN, M. (eds.) *The Semantic Web – ISWC 2005.* Springer Berlin / Heidelberg.

GRUDIN, J. Year. Why CSCW applications fail: problems in the design and evaluationof of organizational interfaces. *In:* ACM conference on Computer-supported cooperative work, 1988 Portland, Oregon, United States. ACM, 85-93.

GRUDIN, J. 1994. Computer-supported cooperative work: history and focus. 27. Available: http://research.microsoft.com/en-us/um/redmond/groups/coet/grudin/papers/ieeecomputer1994.pdf [Accessed 12.03.2011].

HEILFERTY, C. M. 2009. Toward a theory of online communication in illness: concept analysis of illness blogs. *Journal of Advanced Nursing,* 65, 1539-1547.

HJORT, P. F. & WAALER, H. T. 2010. Demens frem mot 2050. *Tidsskrift for den norske legeforening*, 1356-1358.

IAVINDRASANA, J., COHEN, G., DEPEURSINGE, A., MÜLLER, H., MEYER, R. & GEISSBUHLER, A. 2009. Clinical data mining: a review. *Yearbook of medical informatics*, 121-133.

ISO 1999. Human-centred design processes for interactive systems (ISO 13407:1999). Brussels.

KLIMOV, D., SHAHAR, Y. & TAIEB-MAIMON, M. 2010. Intelligent visualization and exploration of time-oriented data of multiple patients. *Artificial Intelligence in Medicine,* 49, 11-31.

KUHLTHAU, C. 1999. Accommodating the User's Information Search Process: Challenges for Information Retrieval System Designers. *Bul. Am. Soc. Info. Sci. Tech.,* 25, 12-16.

MAGNUS, P., ARNESEN, E., HOLMEN, J., STOLTENBERG, C., SØGAARD, A. J. & TELL, G. S. 2003. CONOR – COhort NORway: historie, formål og potensiale. *Norsk Epidemiologi,* 13, 79-82.

MARCHIONINI, G. 1997. *Information Seeking in Electronic Environments (Cambridge Series on Human-Computer Interaction)*, Cambridge University Press.

MARSHALL, W. & HALEY, R. 2000. Use of a Secure Internet Web Site for Collaborative Medical Research. *JAMA: The Journal of the American Medical Association,* 284, 1843-1849.

MORALES RODRIGUEZ, M., CASPER, G. & BRENNAN, P. F. 2007. Patient-centered design. The potential of user-centered design in personal health records. *Journal of AHIMA / American Health Information Management Association,* 78.

MYERS, M. & NEWMAN, M. 2007. The qualitative interview in IS research: Examining the craft. *Information and Organization,* 17, 2-26.

OATES, B. J. 2006. *Researching Information Systems and Computing*, SAGE Publications Ltd.

PLAISANT, C., MUSHLIN, R., SNYDER, A., LI, J., HELLER, D. & SHNEIDERMAN, B. 1998. LifeLines: using visualization to enhance navigation and analysis of patient records. *AMIA American Medical Informatic Association Annual Fall Symposium.* Orlando.

PRATT, W. 2004. Incorporating ideas from computer-supported cooperative work. *Journal of Biomedical Informatics,* 37, 128-137.

PRATT, W., REDDY, M., MCDONALD, D., TARCZY-HORNOCH, P. & GENNARI, J. 2004. Incorporating ideas from computer-supported cooperative work. *Journal of Biomedical Informatics,* 37, 128-137.

ROBSON, C. 2002. *Real World Research*, Blackwell Publishing Limited.

SAFRAN, C., BLOOMROSEN, M., HAMMOND, W., LABKOFF, S., MARKEL-FOX, S., TANG, P., DETMER, D. & WITH INPUT FROM THE EXPERT PANEL SEE APPENDIX, A. 2007. Toward a National Framework for the Secondary Use of Health Data: An American Medical Informatics Association White Paper. *Journal of the American Medical Informatics Association,* 14, 1-9.

SHAHAR, Y., GOREN-BAR, D., GALPERIN-AIZENBERG, M., BOAZ, D. & TAHAN, G. 2003. KNAVE-II: A Distributed Architecture for Interactive Visualization and Intelligent Exploration of Time-Oriented Clinical Data. *Intelligent Data Analysis in Medicine and Pharmacology*

SHNEIDERMAN, B. 1996. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *IEEE Symposium on Visual Languages.* Boulder, Colorado.

SHNEIDERMAN, B. 2002. Inventing discovery tools: combining information visualization with data mining. *Information Visualization,* 1, 5-12.

SPENCE, R. 2007. *Information Visualization: Design for Interaction (2nd Edition)*, Prentice Hall.

STARK, K., SCHULTE, J., HAMPEL, T., SCHIKUTA, E., ZATLOUKAL, K. & EDER, J. 2008. GATiB-CSCW, Medical Research Supported by a Service-Oriented Collaborative System. *In:* BELLAHSÈNE, Z. & LÉONARD, M. (eds.) *Advanced Information Systems Engineering.* Springer Berlin / Heidelberg.

TUFTE, E. 1997. *Visual Explanations: Images and Quantities, Evidence and Narrative,* Cheshire, Connecticut, Graphics Press.

VAN DE BELT, T., ENGELEN, L., BERBEN, S. & SCHOONHOVEN, L. 2010. Definition of Health 2.0 and Medicine 2.0: a systematic review. *Journal of Medical Internet Research,* 12.

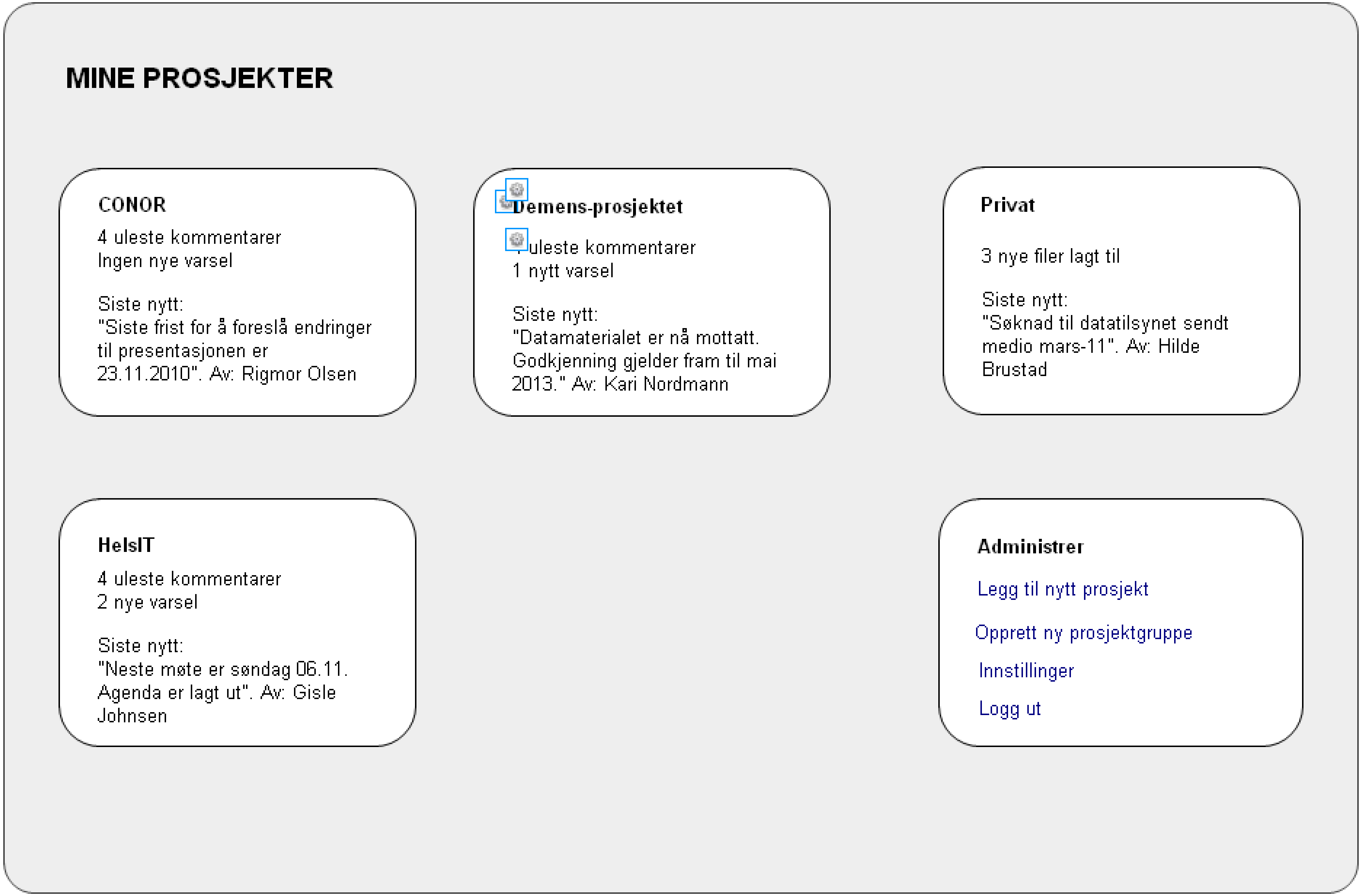
WANG, T., PLAISANT, C., QUINN, A., STANCHAK, R., MURPHY, S. & SHNEIDERMAN, B. Year. Aligning temporal data by sentinel events: discovering patterns in electronic health records. *In:* CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, 2008 Florence, Italy. ACM, 457-466.

WANG, T., PLAISANT, C. & SHNEIDERMAN, B. Year. Visual Information Seeking in Multiple Electronic Health Records: Design Recommendations and A Process Model. *In:* ACM International Health Informatics Symposium 2010 New York, USA.

WONGSUPHASAWAT, K., JOHN ALEXIS GUERRA, G., PLAISANT, C., DAVIDWANG, T., SHNEIDERMAN, B. & TAIEB-MAIMON, M. Year. LifeFlow: Visualizing an Overview of Event Sequences. *In:* ACM CHI Conference on Human Factors in Computing Systems 2011 Vancouver.

Appendix A : Prototype

In the following pages, all the screenshots presented to the researchers are provided.



MINE PROSJEKTER

CONOR

4 uleste kommentarer
Ingen nye varsel

Siste nytt:
"Siste frist for å foreslå endringer til presentasjonen er 23.11.2010". Av: Rigmor Olsen

Demens-prosjektet

uleste kommentarer
1 nytt varsel

Siste nytt:
"Datamaterialet er nå mottatt. Godkjenning gjelder fram til mai 2013." Av: Kari Nordmann

Privat

3 nye filer lagt til

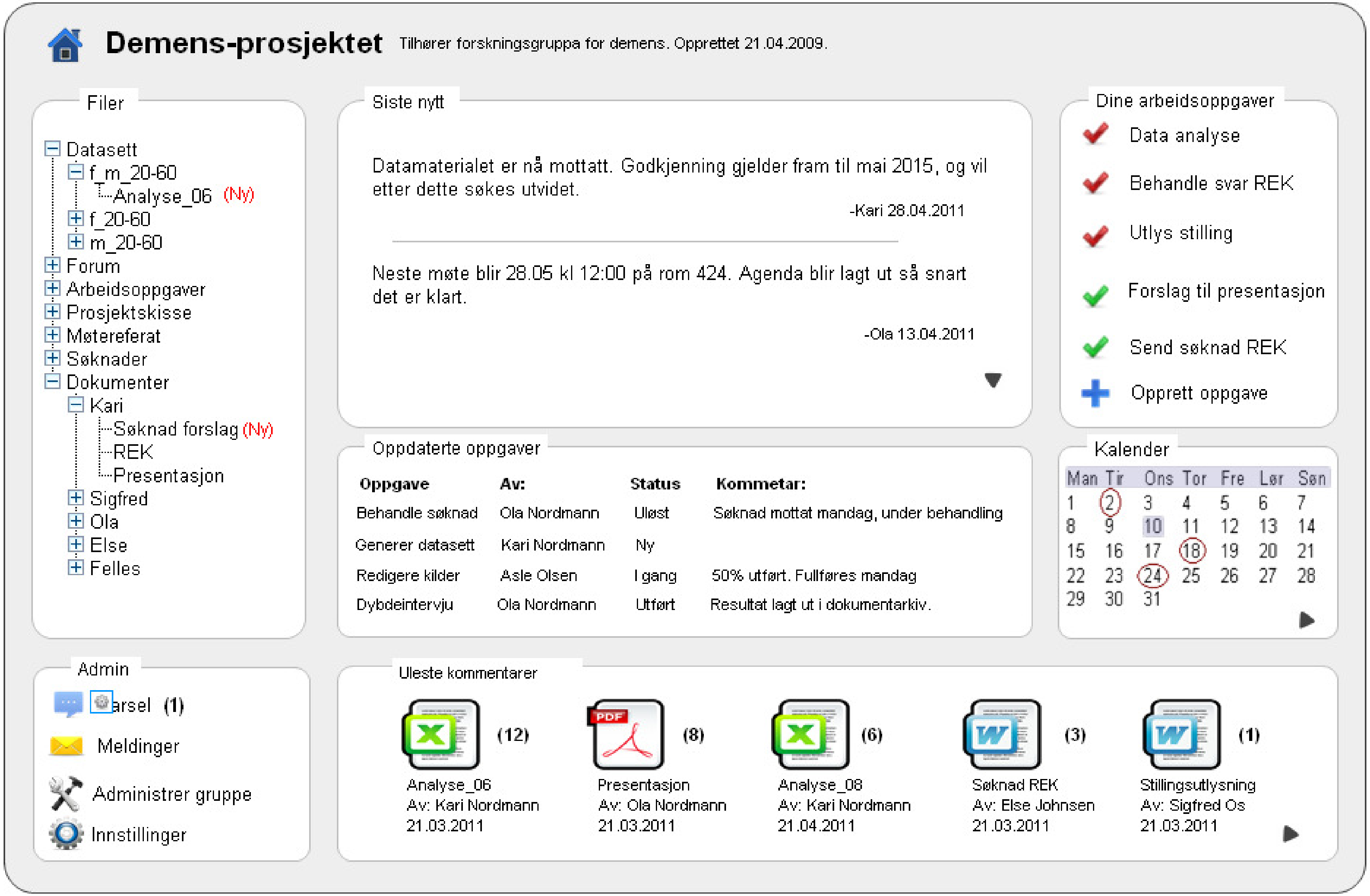
Siste nytt:
"Søknad til datatilsynet sendt medio mars-11". Av: Hilde Brustad

HelsIT

4 uleste kommentarer
2 nye varsel

Siste nytt:
"Neste møte er søndag 06.11. Agenda er lagt ut". Av: Gisle Johnsen

Administrer



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

Filer

- ⊟ Datasett
  - ⊟ f_m_20-60
    - Analyse_06 (Ny)
  - ⊞ f_20-60
  - ⊞ m_20-60
- ⊞ Forum
- ⊞ Arbeidsoppgaver
- ⊞ Prosjektskisse
- ⊞ Møtereferat
- ⊞ Søknader
- ⊟ Dokumenter
  - ⊟ Kari
    - Søknad forslag (Ny)
    - REK
    - Presentasjon
  - ⊞ Sigfred
  - ⊞ Ola
  - ⊞ Else
  - ⊞ Felles

Siste nytt

Datamaterialet er nå mottatt. Godkjenning gjelder fram til mai 2015, og vil etter dette søkes utvidet.

-Kari 28.04.2011

Neste møte blir 28.05 kl 12:00 på rom 424. Agenda blir lagt ut så snart det er klart.

-Ola 13.04.2011

Dine arbeidsoppgaver

- ✓ Data analyse
- ✓ Behandle svar REK
- ✓ Utlys stilling
- ✓ Forslag til presentasjon
- ✓ Send søknad REK
- + Opprett oppgave

Oppdaterte oppgaver

Oppgave	Av:	Status	Kommetar:
Behandle søknad	Ola Nordmann	Uløst	Søknad mottat mandag, under behandling
Generer datasett	Kari Nordmann	Ny	
Redigere kilder	Asle Olsen	I gang	50% utført. Fullføres mandag
Dybdeintervju	Ola Nordmann	Utført	Resultat lagt ut i dokumentarkiv.

Kalender

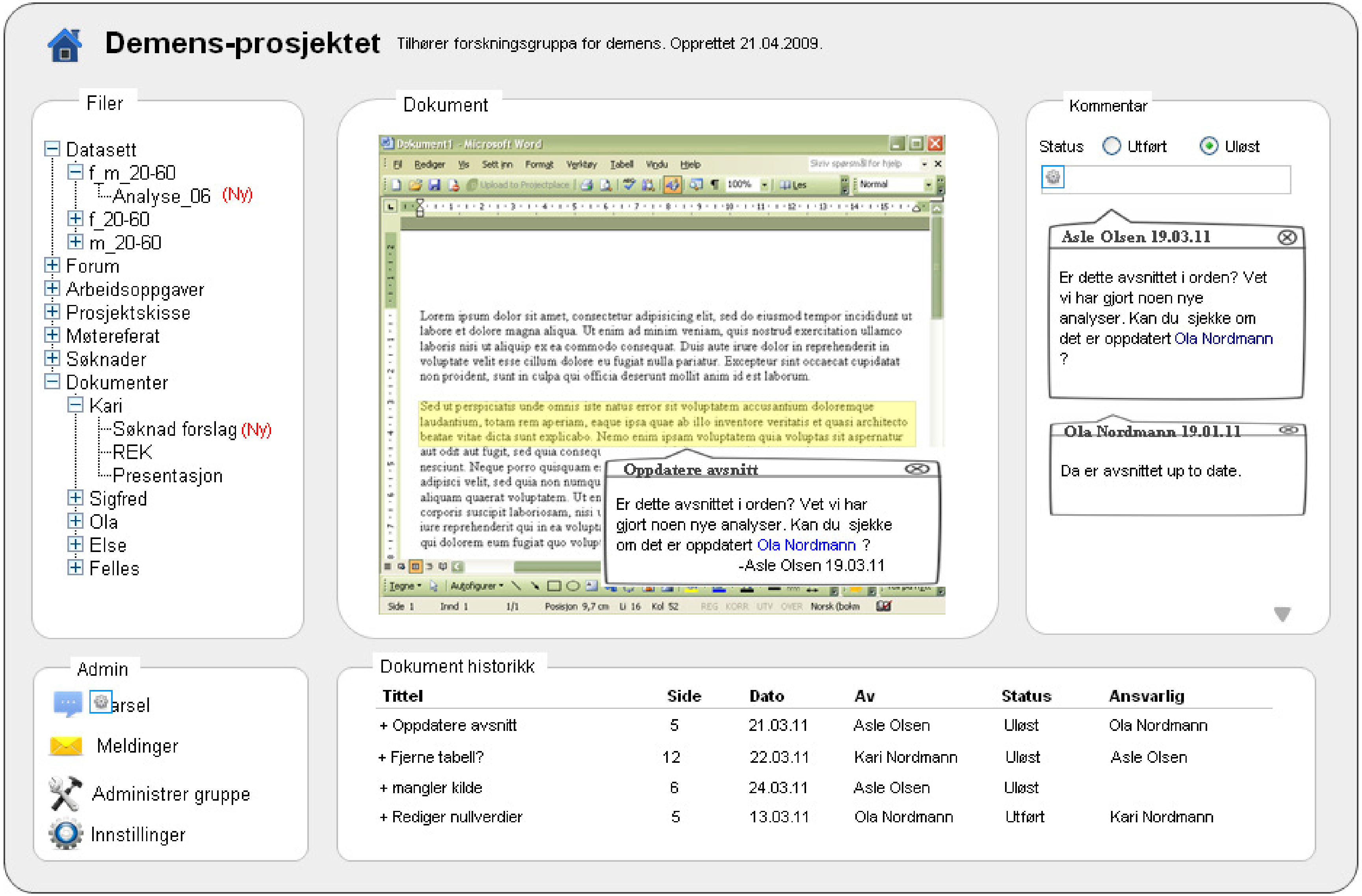
Man	Tir	Ons	Tor	Fre	Lør	Søn
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

Admin

- arsel (1)
- Meldinger
- Administrer gruppe
- Innstillinger

Uleste kommentarer

(12)	(8)	(6)	(3)	(1)
Analyse_06 Av: Kari Nordmann 21.03.2011	Presentasjon Av: Ola Nordmann 21.03.2011	Analyse_08 Av: Kari Nordmann 21.04.2011	Søknad REK Av: Else Johnsen 21.03.2011	Stillingsutlysning Av: Sigfred Os 21.03.2011



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

Filer

- Datasett
  - f_m_20-60
    - Analyse_06 (Ny)
  - f_20-60
  - m_20-60
- Forum
- Arbeidsoppgaver
- Prosjektskisse
- Møtereferat
- Søknader
- Dokumenter
  - Kari
    - Søknad forslag (Ny)
    - REK
    - Presentasjon
  - Sigfred
  - Ola
  - Else
  - Felles

Dokument

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequ nesciunt. Neque porro quisquam e adipisci velit, sed quia non numqu aliquam quaerat voluptatem. Ut en corporis suscipit laboriosam, nisi u iure reprehenderit qui in ea volupt qui dolorem eum fugiat quo volup

Oppdatere avsnitt

Er dette avsnittet i orden? Vet vi har gjort noen nye analyser. Kan du sjekke om det er oppdatert Ola Nordmann ?
-Asle Olsen 19.03.11

Kommentar

Status ○ Utført ● Uløst

Asle Olsen 19.03.11

Er dette avsnittet i orden? Vet vi har gjort noen nye analyser. Kan du sjekke om det er oppdatert Ola Nordmann ?

Ola Nordmann 19.01.11

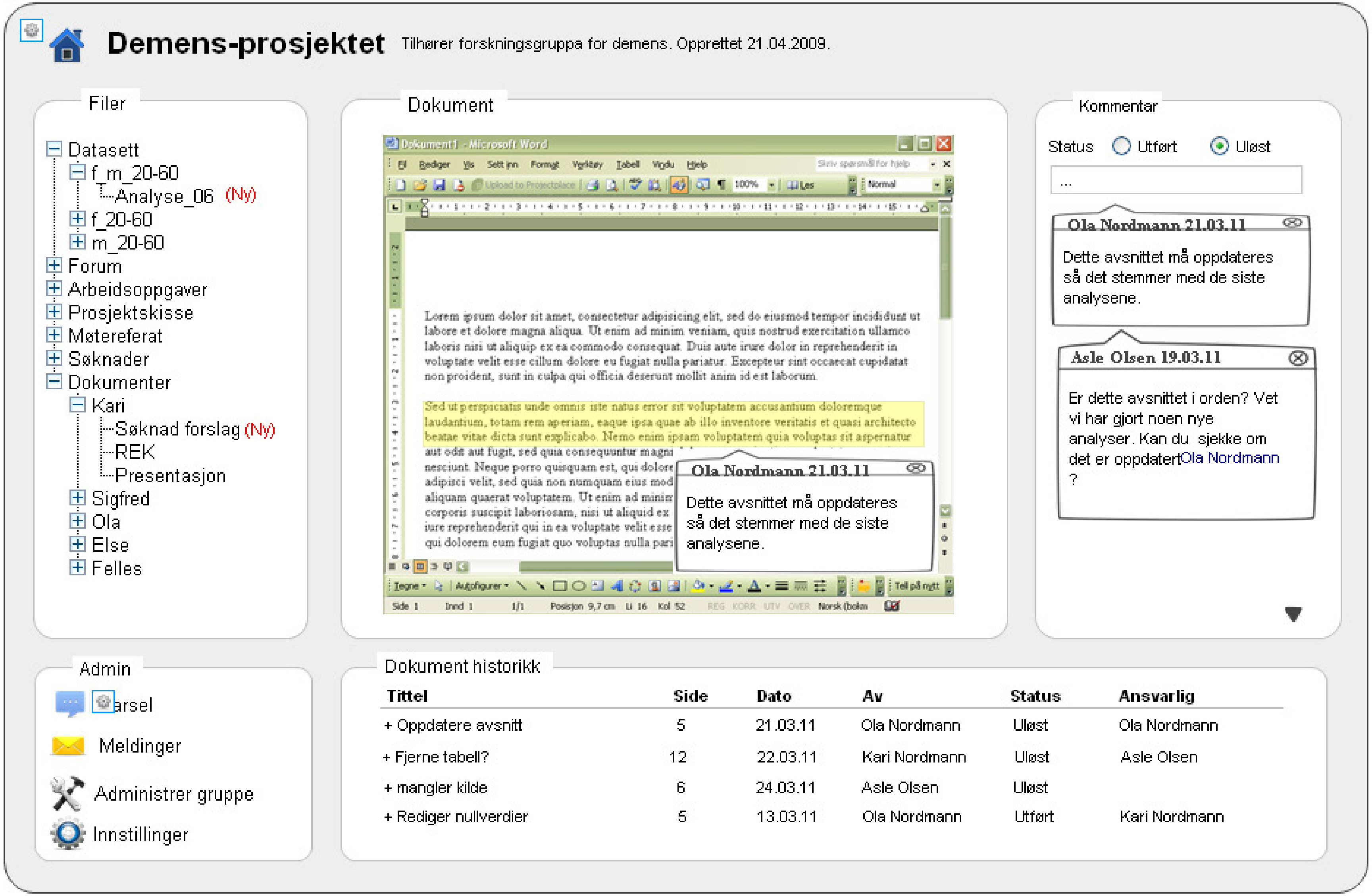
Da er avsnittet up to date.

Admin

- arsel
- Meldinger
- Administrer gruppe
- Innstillinger

Dokument historikk

Tittel	Side	Dato	Av	Status	Ansvarlig
+ Oppdatere avsnitt	5	21.03.11	Asle Olsen	Uløst	Ola Nordmann
+ Fjerne tabell?	12	22.03.11	Kari Nordmann	Uløst	Asle Olsen
+ mangler kilde	6	24.03.11	Asle Olsen	Uløst	
+ Rediger nullverdier	5	13.03.11	Ola Nordmann	Utført	Kari Nordmann



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

Filer

- ⊟ Datasett
  - ⊟ f_m_20-60
    - Analyse_06 (Ny)
  - ⊞ f_20-60
  - ⊞ m_20-60
- ⊞ Forum
- ⊞ Arbeidsoppgaver
- ⊞ Prosjektskisse
- ⊞ Møtereferat
- ⊞ Søknader
- ⊟ Dokumenter
  - ⊟ Kari
    - Søknad forslag (Ny)
    - REK
    - Presentasjon
  - ⊞ Sigfred
  - ⊞ Ola
  - ⊞ Else
  - ⊞ Felles

Dokument

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni nesciunt. Neque porro quisquam est, qui dolore adipisci velit, sed quia non numquam eius mod aliquam quaerat voluptatem. Ut enim ad minim corporis suscipit laboriosam, nisi ut aliquid ex iure reprehenderit qui in ea voluptate velit esse qui dolorem eum fugiat quo voluptas nulla pari

Ola Nordmann 21.03.11

Dette avsnittet må oppdateres så det stemmer med de siste analysene.

Kommentar

Status ○ Utført ● Uløst

...

Ola Nordmann 21.03.11

Dette avsnittet må oppdateres så det stemmer med de siste analysene.

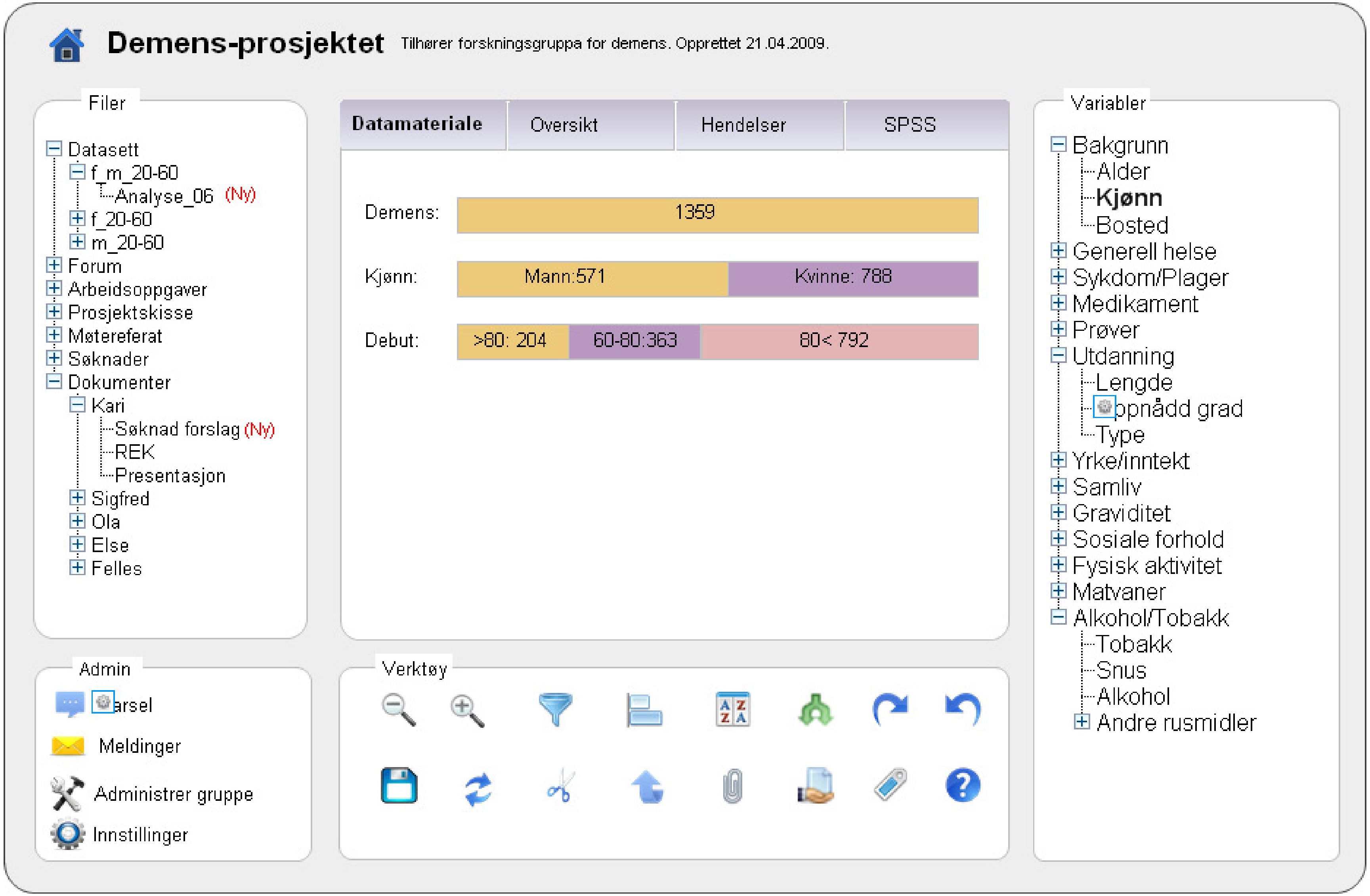
Asle Olsen 19.03.11

Er dette avsnittet i orden? Vet vi har gjort noen nye analyser. Kan du sjekke om det er oppdatert Ola Nordmann ?

Admin

Dokument historikk

Tittel	Side	Dato	Av	Status	Ansvarlig
+ Oppdatere avsnitt	5	21.03.11	Ola Nordmann	Uløst	Ola Nordmann
+ Fjerne tabell?	12	22.03.11	Kari Nordmann	Uløst	Asle Olsen
+ mangler kilde	6	24.03.11	Asle Olsen	Uløst	
+ Rediger nullverdier	5	13.03.11	Ola Nordmann	Utført	Kari Nordmann

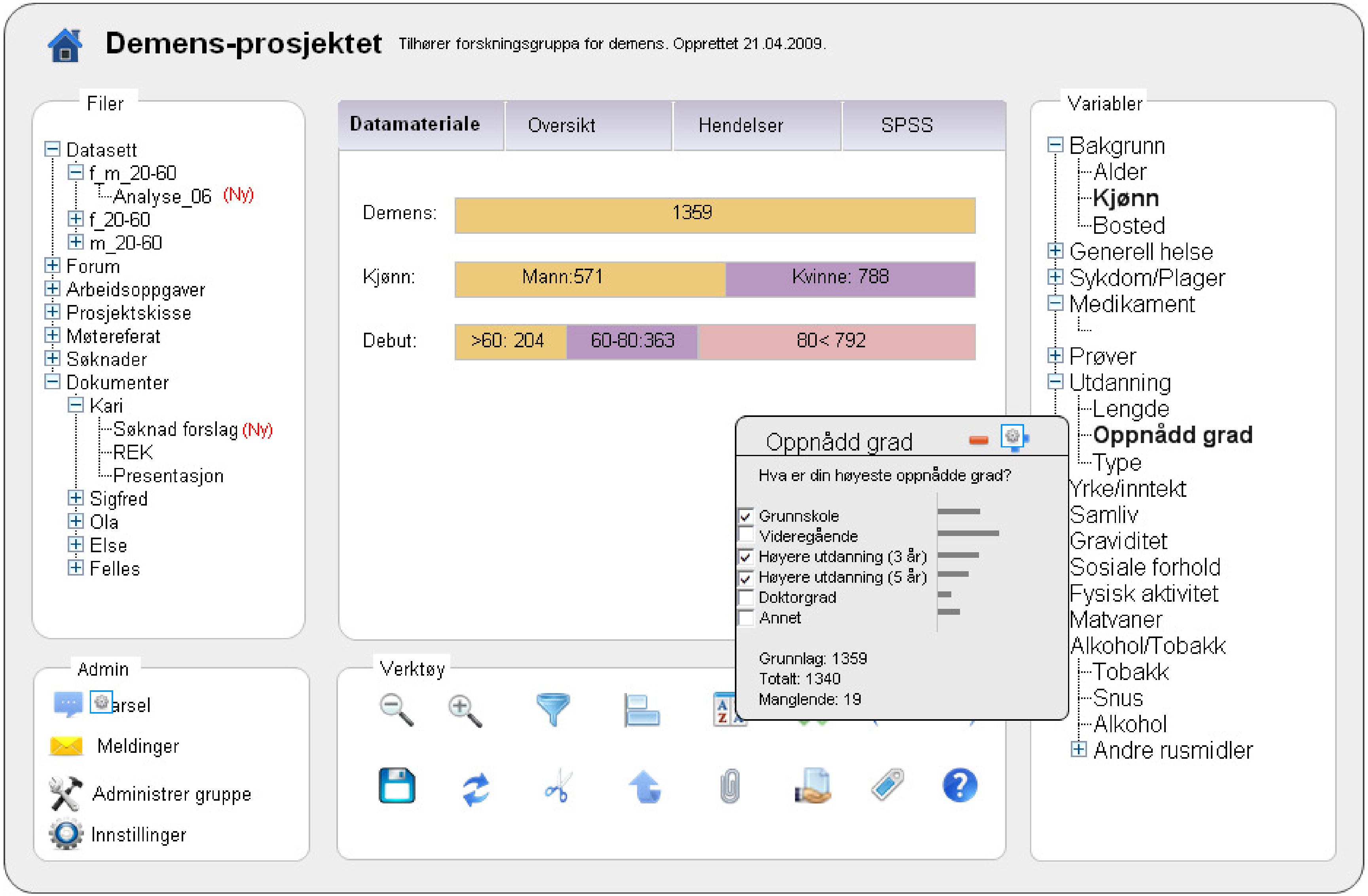


Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

Demens: 1359

Kjønn: Mann:571 Kvinne: 788

Debut: >80: 204 60-80:363 80< 792



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

Filer

- Datasett
  - f_m_20-60
    - Analyse_06 (Ny)
  - f_20-60
  - m_20-60
- Forum
- Arbeidsoppgaver
- Prosjektskisse
- Møtereferat
- Søknader
- Dokumenter
  - Kari
    - Søknad forslag (Ny)
    - REK
    - Presentasjon
  - Sigfred
  - Ola
  - Else
  - Felles

Datamateriale | Oversikt | Hendelser | SPSS

Demens: 1359

Kjønn: Mann:571 | Kvinne: 788

Debut: >60: 204 | 60-80:363 | 80< 792

Oppnådd grad

Hva er din høyeste oppnådde grad?

☑ Grunnskole
☐ Videregående
☑ Høyere utdanning (3 år)
☑ Høyere utdanning (5 år)
☐ Doktorgrad
☐ Annet

Grunnlag: 1359
Totalt: 1340
Manglende: 19

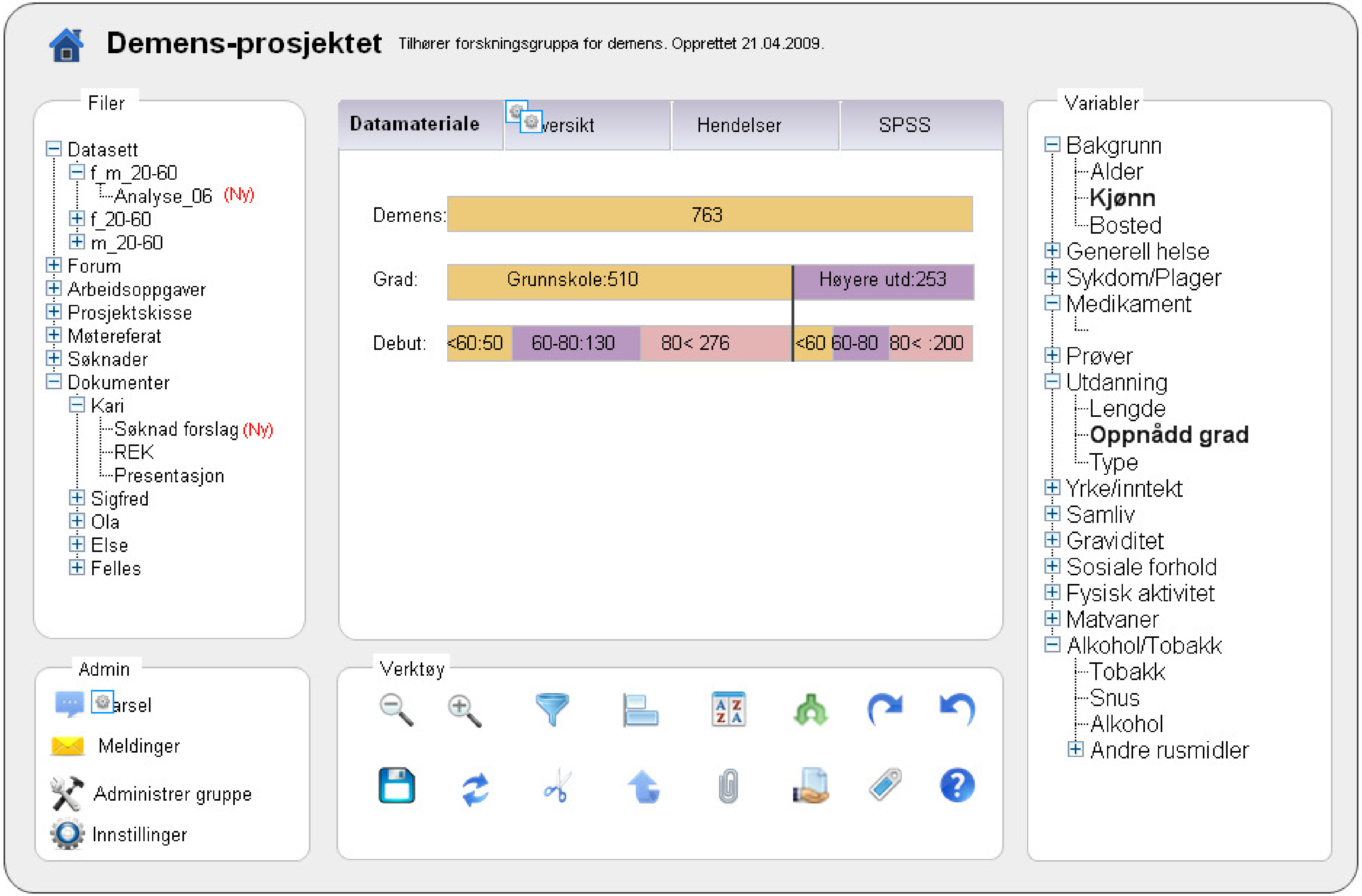
Variabler

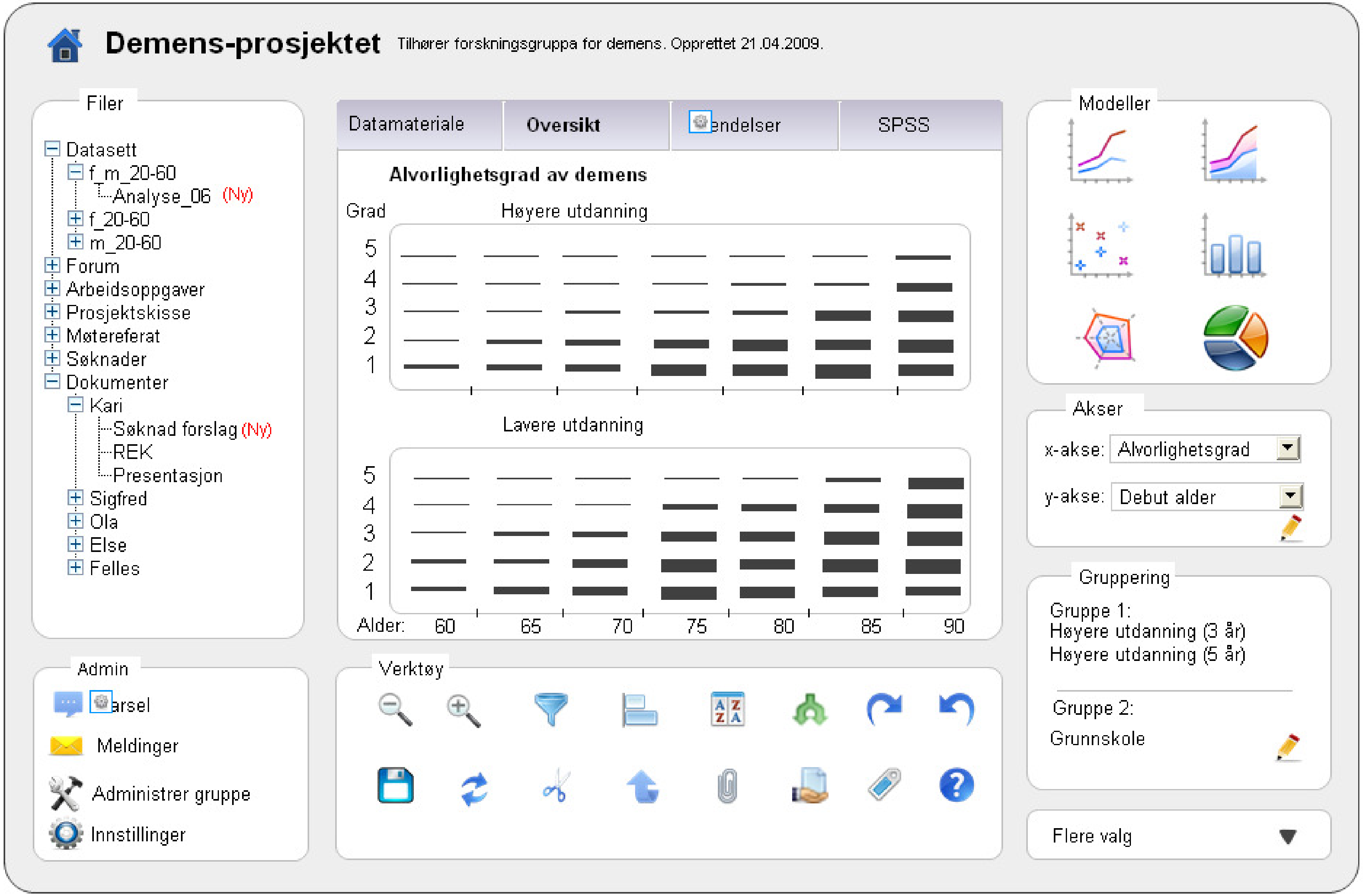
- Bakgrunn
  - Alder
  - **Kjønn**
  - Bosted
- Generell helse
- Sykdom/Plager
- Medikament
- Prøver
- Utdanning
  - Lengde
  - **Oppnådd grad**
  - Type
- Yrke/inntekt
- Samliv
- Graviditet
- Sosiale forhold
- Fysisk aktivitet
- Matvaner
- Alkohol/Tobakk
  - Tobakk
  - Snus
  - Alkohol
  - Andre rusmidler

Admin

- Varsel
- Meldinger
- Administrer gruppe
- Innstillinger

Verktøy


Demens-prosjektet
Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.
Filer
Datasett
f_m_20-60
Analyse_06 (Ny)
f_20-60
m_20-60
Forum
Arbeidsoppgaver
Prosjektskisse
Møtereferat
Søknader
Dokumenter
Kari
Søknad forslag (Ny)
REK
Presentasjon
Sigfred
Ola
Else
Felles
Datamateriale
versikt
Hendelser
SPSS
Demens:
763
Grad:
Grunnskole:510
Høyere utd:253
Debut:
<60:50
60-80:130
80< 276
<60
60-80
80< :200
Variabler
Bakgrunn
Alder
Kjønn
Bosted
Generell helse
Sykdom/Plager
Medikament
Prøver
Utdanning
Lengde
Oppnådd grad
Type
Yrke/inntekt
Samliv
Graviditet
Sosiale forhold
Fysisk aktivitet
Matvaner
Alkohol/Tobakk
Tobakk
Snus
Alkohol
Andre rusmidler
Admin
arsel
Meldinger
Administrer gruppe
Innstillinger
Verktøy



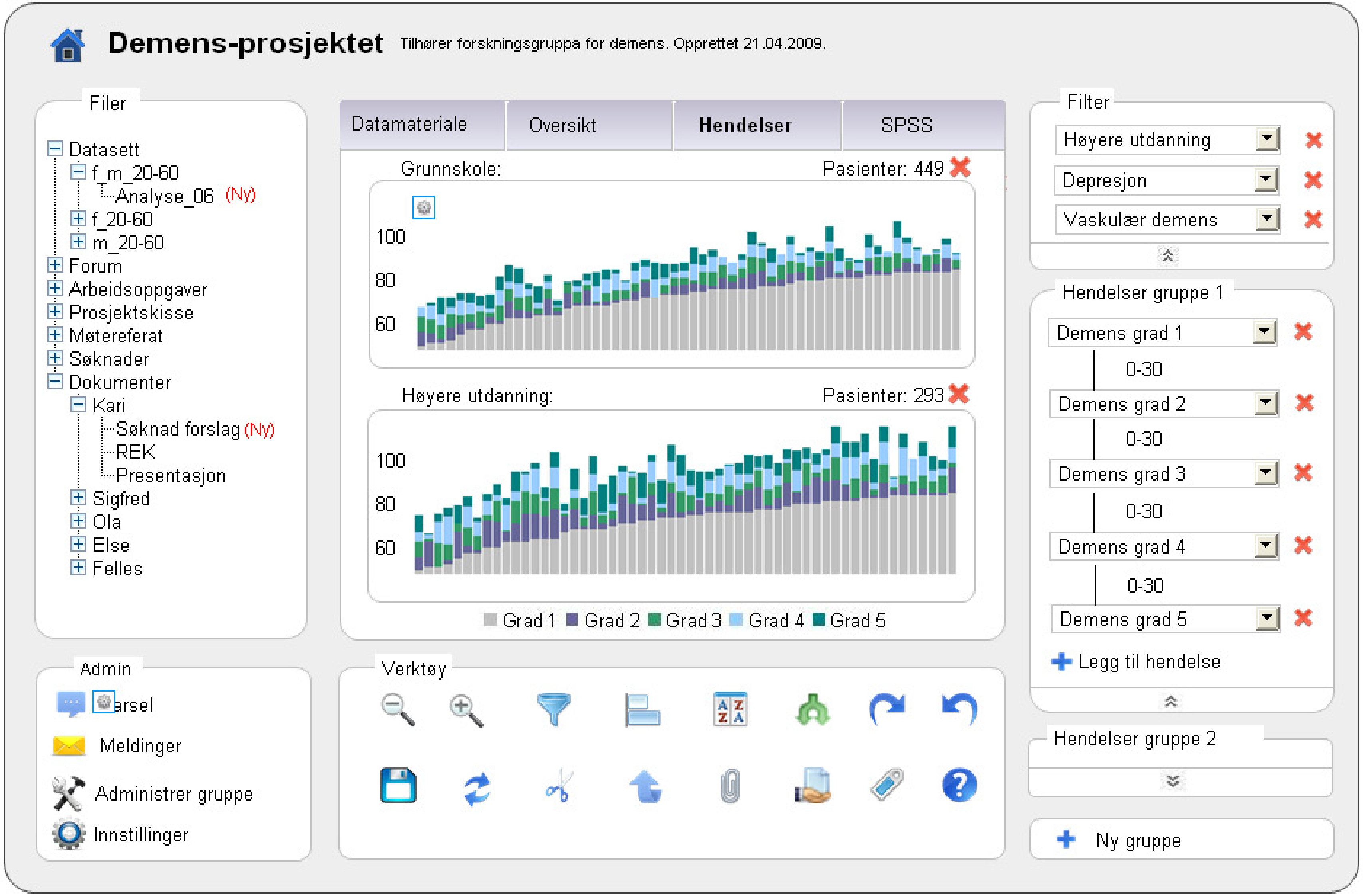
Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.

x-akse: Alvorlighetsgrad

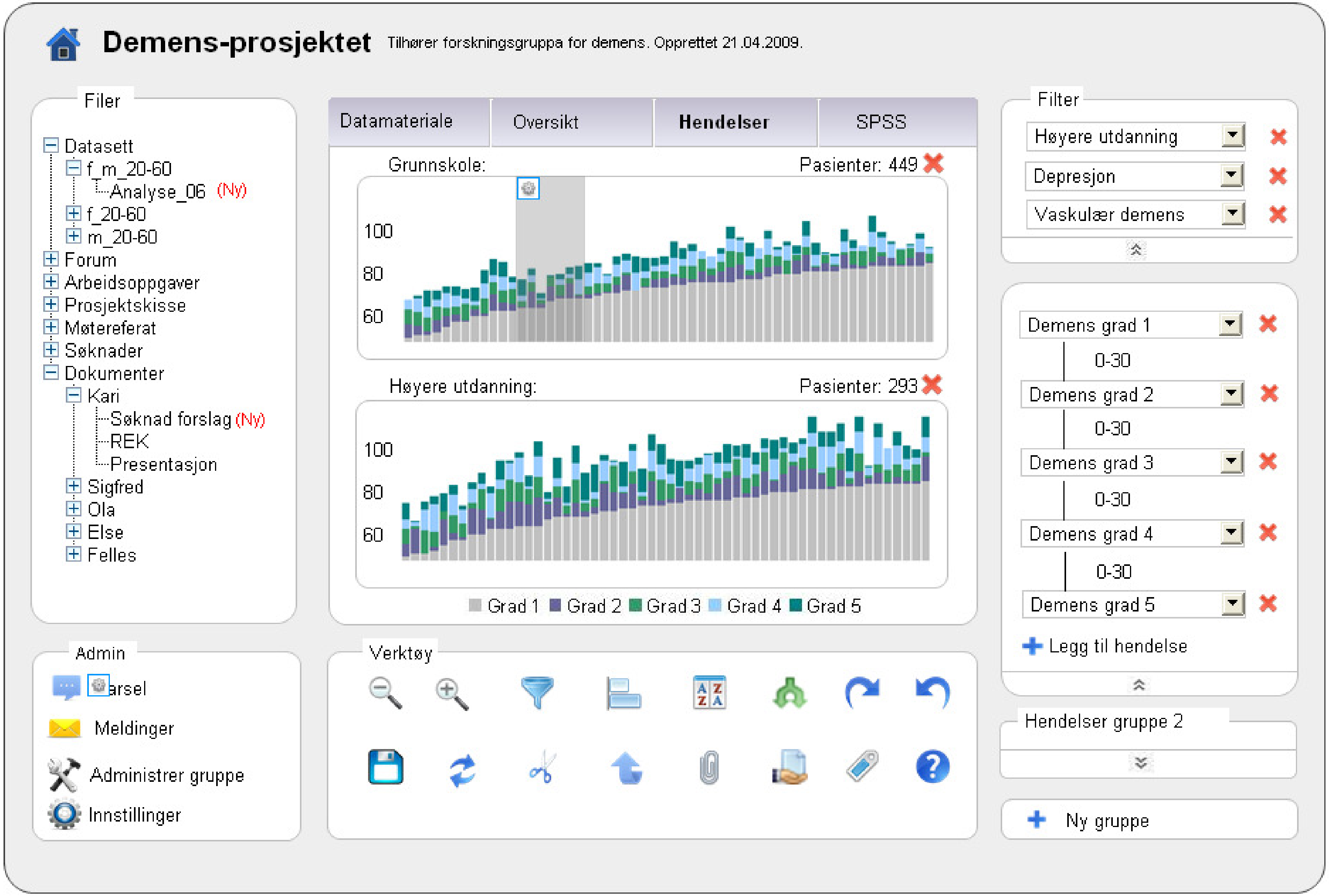
y-akse: Debut alder

Gruppe 1:
Høyere utdanning (3 år)
Høyere utdanning (5 år)

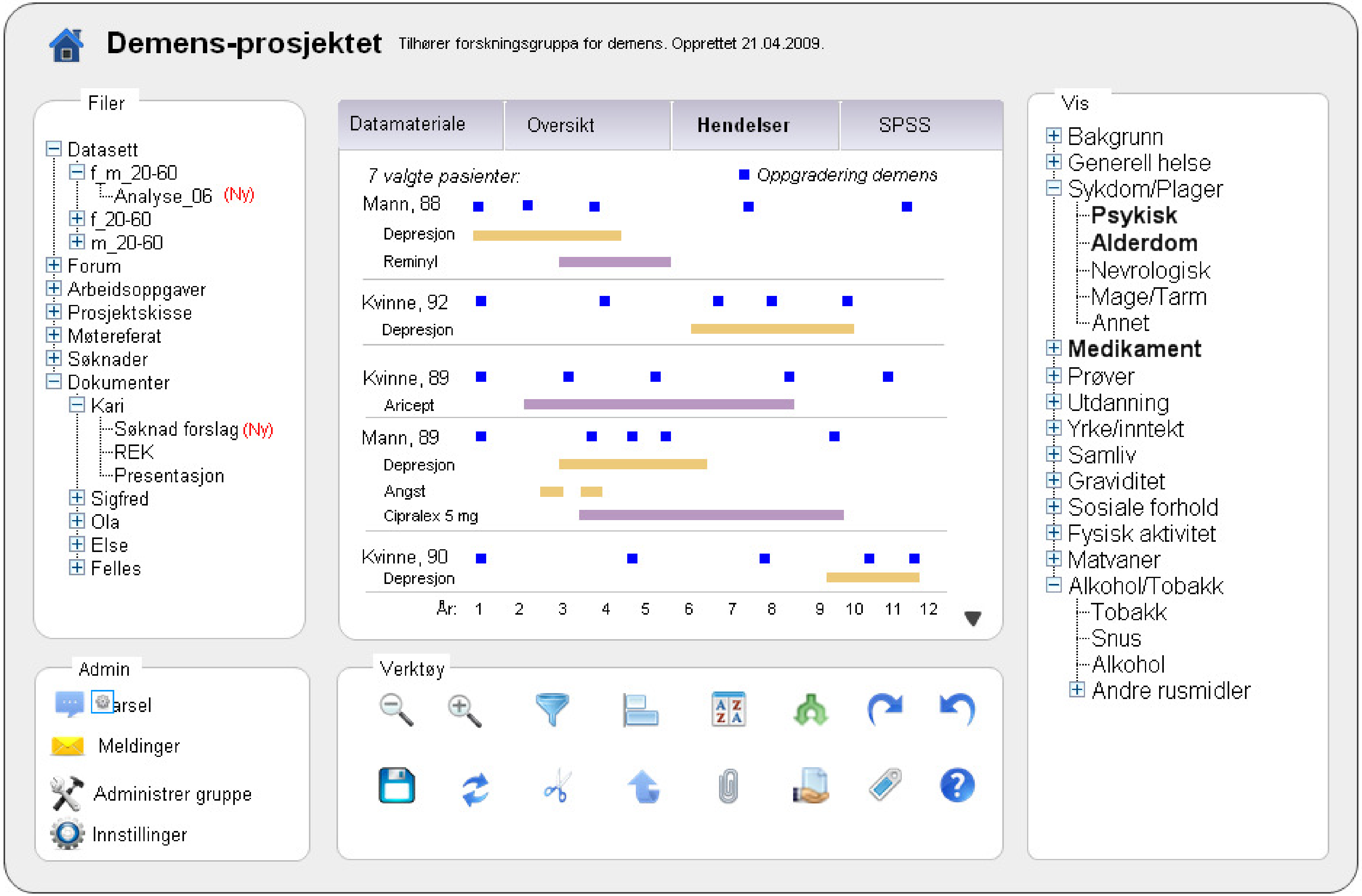
Gruppe 2:
Grunnskole



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.



Demens-prosjektet Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.


Demens-prosjektet
Tilhører forskningsgruppa for demens. Opprettet 21.04.2009.
Filer
Datasett
f_m_20-60
Analyse_06 (Ny)
f_20-60
m_20-60
Forum
Arbeidsoppgaver
Prosjektskisse
Møtereferat
Søknader
Dokumenter
Kari
Søknad forslag (Ny)
REK
Presentasjon
Sigfred
Ola
Else
Felles
Datamateriale
Oversikt
Hendelser
SPSS
7 valgte pasienter:
Oppgradering demens
Mann, 88
Depresjon
Reminyl
Kvinne, 92
Depresjon
Kvinne, 89
Aricept
Mann, 89
Depresjon
Angst
Cipralex 5 mg
Kvinne, 90
Depresjon
År: 1 2 3 4 5 6 7 8 9 10 11 12
Vis
Bakgrunn
Generell helse
Sykdom/Plager
Psykisk
Alderdom
Nevrologisk
Mage/Tarm
Annet
Medikament
Prøver
Utdanning
Yrke/inntekt
Samliv
Graviditet
Sosiale forhold
Fysisk aktivitet
Matvaner
Alkohol/Tobakk
Tobakk
Snus
Alkohol
Andre rusmidler
Admin
arsel
Meldinger
Administrer gruppe
Innstillinger
Verktøy