University of Bergen

Department of Information Science and Media Studies

Master Thesis

Arthroplasty Data Visualization

Author: **Arle Farsund Solheim**

Supervisor: **Dr. Ankica Babic**

June 14, 2021

“Things are only impossible until they are not.”

Cpt. Jean-Luc Picard

Abstract

This master's thesis presents the work done in the field of visualization and interactivity conducted within the Design Science framework. The main goal was to make the data analysis using the arthroplasty register data into a more independent, easy, and user-friendly experience. The visualization artifact was created to support presentation of data material and results from data mining with a purpose of understand patient outcomes, longevity of implants, and present demographic and other data in a more contemporary way. There is a wealth of information and reports at the website of the Norwegian Arthroplasty Register, but very little in terms of interactivity and independent user exploration of data.

The work was carried out as a part of a back- and front-end development with data mining methods developed for knee and hip prosthesis data being the back-end, and the front-end consisted of a user interface in addition to visualization. This setup had several advantages, where the selection of data mining methods and implementation of a high-fidelity user interface all contributed to a better user experience of the visualizations.

The resulting artifact is comprised of visualizations of demographic data, Kaplan-Meier, and an interactive map of Norway. Interactivity enabled exploring data for selected periods of time, comparison of performance in different prostheses, and exploring patient population behind certain points on a survival graph. The map of Norway offers features such as demographic data and comparison of top 5 prostheses in different counties.

The evaluation was carried out with the use of three different evaluation tools and interviews with domain and usability experts. Feedback during interviews was encouraging and indicated the potential usefulness of the visualizations.

The system in its current form is more directed towards expert users, but can be easily adjusted to patients and the wider public, which could be a subject of future research. More visualizations and data analytical methods could further enhance the current solutions.

Acknowledgements

I would like to firstly thank my supervisor Dr. Ankica Babic, whose inspiration, motivation, input, support, and guidance has been crucial in the realization of this thesis. I honestly could not have asked for a better supervisor.

I also extend my gratitude to Dr. Peter Ellison, who offered valuable time, help, and feedback to our team. The Norwegian Arthroplasty Register has my thanks for the foundation and data this thesis is based on.

Thank you to my friend, neighbor, and fellow student Are Nyhammer for talks, walks, and motivation during this weird year.

Last but not least a big thank you to my teammates Knut, Sunniva, and Sølve for keeping me going, lightening an otherwise unbearable load, and elevating my research.

Contents

List of Figures

List of Tables

Abbreviations

NAR - Norwegian Arthroplasty Register

HCI - Human Computer Interaction

SUS - System Usability Scale

SVG - Scalable Vector Graphics

API - Application Programming Interface

RTD - Research Through Design

1 Introduction

The basis for this thesis is a cooperation with the Norwegian Arthroplasty Register. The register was started in 1987 and contains data about hip, knee, shoulder, and other joint replacement surgeries, i.e. arthroplasty, from all Norwegian hospitals and is routinely updated. The register also contains demographics about patients, as well as data about patient survival, comorbidity, pain, function, outcome satisfaction etc. [21].

The data in the register contains sensitive medical data about real patients, and is therefore not freely available to the public. The register presents research, results, and anonymized data in an annual report that is released publicly. These reports present data in tables and standardized medical graph techniques. While these presentations are established and useful for researchers in the field, this thesis seeks to explore possible improvements in visualizing data from the register for experienced researchers, but also for patients and non-expert user groups. There are several technological developments in the field of visualization and data analysis which could be well utilized to enhance user-experience and enable users to run analyses independently. Software packages for visualization and analysis have made their way into the health domains. Health 2.0, a term suggesting contemporary strategies that are Web-based, participatory, and mobile, provide a useful framework for all kinds of data management. These strategies rely on patient-records data, patient self-monitoring data, and even social media, thanks to which data could be visualized and analysed [47].

Data mining is the process of discovering patterns in large data sets using a combination of machine learning, statistics, and database systems [6]. Patterns found from data mining the register data could give insights into prosthesis survivability, patient quality of life, risk of revision surgery and much more. Using data mining in combination with visualization even allows for use of older well-established multivariate statistical methods such as survival analysis with Kaplan-Meier [30, 45].

Visualizing data using graphs, illustrations, and interactive graphics helps humans to better understand data [1, 28]. There are several user groups with a potential to benefit from visualizations, with doctors and patients being those that are most likely to gain a deeper understanding of facts, challenges, risks, and final outcomes. This thesis focuses on visualization in the context of developing data mining routines for knee and hip arthroplasty as well as designing a Human Computer Interaction interface that will allow users greater independence and a more user friendly environment for exploration of data. The work is part of a collaborative project of four master students of whom two are back-end developers (data mining for knee and hip arthroplasty) and two front-end developers (visualization and an HCI interface). It is in this context that development will be presented and reflected on.

1.1 Research Questions

The research in this thesis will attempt to answer the following questions:

RQ1: How can data visualization help doctors and patients get a better understanding of medical data?

RQ2: Can arthroplasty data from the register be visualized in a more efficient and informative way to meet different user groups' needs?

RQ3: Should sensitive medical data be presented to patients and be open to the public?

1.2 Outline of Thesis

Chapter 2: Literature Review presents theory, literature, and related work relevant to the research.

Chapter 3: Methodology and Methods introduces the methodologies, methods and tools used in the research, design, development, and evaluation during the project.

Chapter 4: Requirements explains the different requirements and ethical considerations for the project.

Chapter 5: Visualization Development details the iterations of development producing the visualizations.

Chapter 6: Results demonstrates the final visualizations and their functionality.

Chapter 7: Evaluation summarizes the evaluation process and the resulting feedback.

Chapter 8: Discussion discusses the research, evaluation, and answers the research questions.

Chapter 9: Conclusions and Future Work sums up the research and gives recommendations for future work.

2 Literature Review

The literature presented in this chapter gives an introduction to the topics relevant to this thesis. It will cover concepts of total joint replacement surgery(arthroplasty), data registers, data mining, data visualization, and patient quality-of-life studies.

2.1 Arthroplasty

Arthroplasty is a collective term for surgical procedures that restore the function of a joint by either resurfacing the bone or replacing it with an artificial joint (prosthesis) [37]. The reason for arthroplasty ranges from trauma, wear and tear through exercise, to medical conditions such as osteoarthritis, a degenerative joint disease that affects the cartilage and cushioning surrounding the joint. The majority of arthroplasty operations are for hip and knee joints, whereas shoulder, ankle, elbow, and fingers are less common [37].

After surgery patients need to go through physical therapy in order to recover mobility in the joint and surrounding muscle. Pain is managed through medicine during this process in order to be able to exercise the joint [37]. This process starts in the hospital but needs to continue after discharge in order to regain muscle strength and a good range of motion [37]. Figure 2.1 shows different types of joint prostheses.

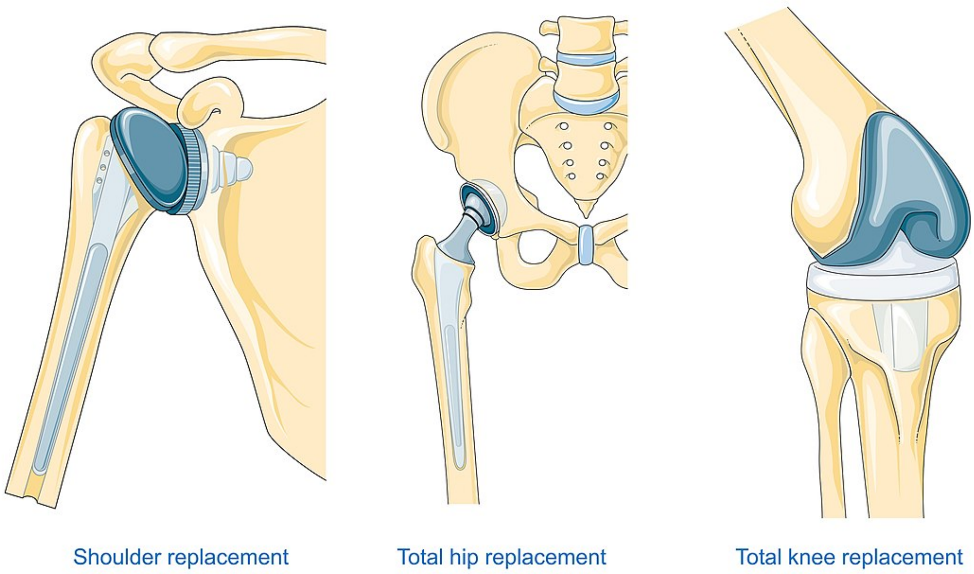


Figure 2.1: Arthroplasty

2.2 Health Data Registries

2.2.1 Categorizing the World of Registries

The term registry, or register, is widely used to refer to any database storing clinical information collected as a byproduct of patient care [13]. Databases are important tools widely used in science, including modern medical practice and research. A database can be defined as a structured repository of data that allows data collection, modification, and retrieval. A register, then, is a more sophisticated database with clearly defined health and demographic data about patients with specific health characteristics and a defined purpose. While registers have been in use for centuries [13], digital technologies have increased their frequency and efficiency over the last few decades.

2.2.2 The Norwegian Arthroplasty Register

In Norway, approximately 8500 hip arthroplasties, 6000 knee arthroplasties, 700 shoulder arthroplasties, and 350 other joint arthroplasties are performed every year, excluding surgeries where old or damaged prostheses are replaced [20].

The register was started as a way to monitor the quality of joint prostheses and discover faulty arthroplasty [20]. Before the register was started there were rising concerns about poor quality prostheses being used in patients without knowing they were poor. While the register keeps track of prostheses from different manufacturers, it does not prevent the use of new types of prostheses that have not yet been tested in studies [20].

The main objective the register states is to prevent the use of poor prostheses in patients [20]. And as a secondary objective to provide an overview of the state of products and procedures being used in the field. The results of the register are published in scientific reports, presented at seminars and conferences. Every hospital in Norway receives reports that pertain to both the specific hospital as well as the country as a whole. These reports are also published on the website of the National Service Centre for Medical Quality Registers [20].

2.3 Data mining

2.3.1 Principles of Data Mining

This article gives an introduction to data mining, its main areas of use, and some of its weaknesses. [19] emphasizes two main parts of data mining; model building and pattern discovery. The article defines model building as a summary of datasets as a whole, mentioning regression models, cluster decomposition, and Bayesian networks as tools used for this in modern statistics [19]. Pattern discovery is the more specific look at the data in search of relevant or interesting patterns. Hand goes more in-depth about patterns. He explains how some patterns can occur as a byproduct of the data recording process, creating false positives. Further, he discusses the importance of recognizing these false positives, especially when the data is in a serious field like medicine. Finally, he emphasizes the question of whether the pattern ”matters”, the scientific, or commercial, value needs to be assessed.

The concepts of model building and pattern discovery are useful terms to describe processes in data mining. The descriptions of pattern discovery and their pitfalls is very relevant and useful. It also introduces data cleaning, an important first step in data mining. The article is on the shorter side but serves as a good introduction and baseline for the basics of data mining.

2.3.2 Data Preparation for Data Mining

This article explains the importance of the quality of data used in data mining. and highlights the consequences of low-quality data. It presents methods and techniques for cleaning and preparing data in ways that improve the outcomes of using it for data mining. The article claims much of data mining is being based on a false assumption that the data is of good quality and containing no missing or incorrect values. They present three consequences of working with this assumption:

1. missing useful patterns that are hidden in noisy data
2. low performance
3. poor-quality outputs

Next, the article presents the steps involved in the process of data mining. There are four steps; defining the problem, data pre-processing, data mining, and post data mining. The article then further elaborates on these steps:

Defining the problem - The goals of a knowledge discovery project must be identified. The goals must be verified as actionable. For example, if the goals are met, a business organization can then put the newly discovered knowledge to use. The data to be used must also be identified clearly. [50]

Data pre-processing - Data preparation comprises those techniques concerned with analyzing raw data so as to yield quality data, mainly including data collecting, data integration, data transformation, data cleaning, data reduction, and data discretization. [50]

Data mining - Given the cleaned data, intelligent methods are applied in order to extract data patterns. Patterns of interest are searched for, including classification rules or trees, regression, clustering, sequence modeling, dependency, and so forth. [50]

Post data mining - Post data mining consists of pattern evaluation, deploying the model, maintenance, and knowledge presentation. [50]

They explain that these steps are iterative, in that discoveries made in the data mining process could reveal that additional data cleaning is required.

Further, the article discusses the importance of the preparation of data before beginning data mining and highlights the lack of previous research in this area. The data preparation and cleaning takes up approximately 80% of the total data engineering effort [50]. They present the importance of data preparation through three aspects:

1. Real-world data is impure
2. High-performance mining systems require quality data

3. Quality data yields high-quality patterns.

The consequence of the first aspect is that noisy, incomplete, or inconsistent data can end up disguising useful patterns in the dataset. In this context noisy data is data containing errors or outliers, incomplete data is data lacking attribute values or certain attributes of interest, and inconsistent data is data containing discrepancies in codes or names.
The second aspect points out that cleaning and preparing the data makes the dataset smaller, which improves the efficiency of the data mining process. Tasks in preparing data include filtering and selecting relevant data, removing anomalies, and eliminating duplicates. Reducing the data could also be part of the preparation, selecting the most relevant parts of the dataset to mine.
In the third aspect, they discuss how data preparation leads to quality data, which in turn leads to quality patterns. They give examples of tasks in data preparation that helps with this:

- Recovering incomplete data: filling the values missed, or reducing ambiguity. [50]
- Purifying data: correcting errors, or removing outliers (unusual or exceptional values). [50]
- Resolving data conflicts: using domain knowledge or expert decision to settle discrepancy. [50]

These three aspects show that data pre-processing, cleaning, and preparation is a big part of the data mining process. It also shows how it improves the data mining itself, and is a critical part, although a challenging one.

While this article is on the older side, it does a good job of highlighting the importance of preparing and processing data from a dataset before the process of data mining begins. The four iterative steps presented can be a good baseline to work from. The fourth step even fits well with the data visualization aspect of this project. They reiterate several times in the article that the preparation of the data is the most time-consuming part of the data mining process.

2.3.3 Predictive data mining in clinical medicine: Current issues and guidelines

This paper [3] discusses data mining in the field of clinical medicine and proposes a framework for coping with the problems that arise in this context. The scope of the paper is centered on predictive data mining, i.e. using data mining to try and extrapolate future trends from existing data. The paper opens with an explanation of predictive data mining and its use in the medical field. Using an example of a dataset of hip arthroplasty patients they then present some ways of modeling and presenting the data. There are also some technical descriptions of specific techniques for predictive modeling, as well as the existing standards for predictive data mining at the time.

Bellazzi and Zupan give an overview of the ways predictive data mining already has contributed in the medical field and why it is a method well suited for the field:

> ”Data mining may effectively contribute to the development of clinically useful

> predictive models thanks to at least three inter-related aspects: (a) a comprehensive and purposive approach to data analysis that involves the application of methods and approaches drawn from different scientific areas; (b) the explanatory capability of such models; (c) the capability of using the domain (background) knowledge in the data analysis process. [3]

The paper gives a thorough description of the background for predictive data mining, focusing on its use in clinical medicine. It introduces the difference between descriptive and predictive data mining and discusses some existing standards for these differences. It also touches on aspects of visualizations of data. The paper is very technical, which means it is a great resource for learning about data mining but can be hard to understand without thorough reading.

2.4 Data visualization

Data visualization is the field of graphical representation of data. The aim of visualizing data is usually to represent data in a more understandable way, as humans have an easier time conceptualizing information when analyzing it visually. As a multidisciplinary field visualization uses concepts from statistics, mathematics, HCI, and more as well as using principles from visual arts, as it is a visual medium with potential aesthetic elements. [5] Figure 2.2 is presenting the main areas of development contributing to the data visualization. In short, there are three main fields: Scientific visualization, information visualization, and visual analytics [5].

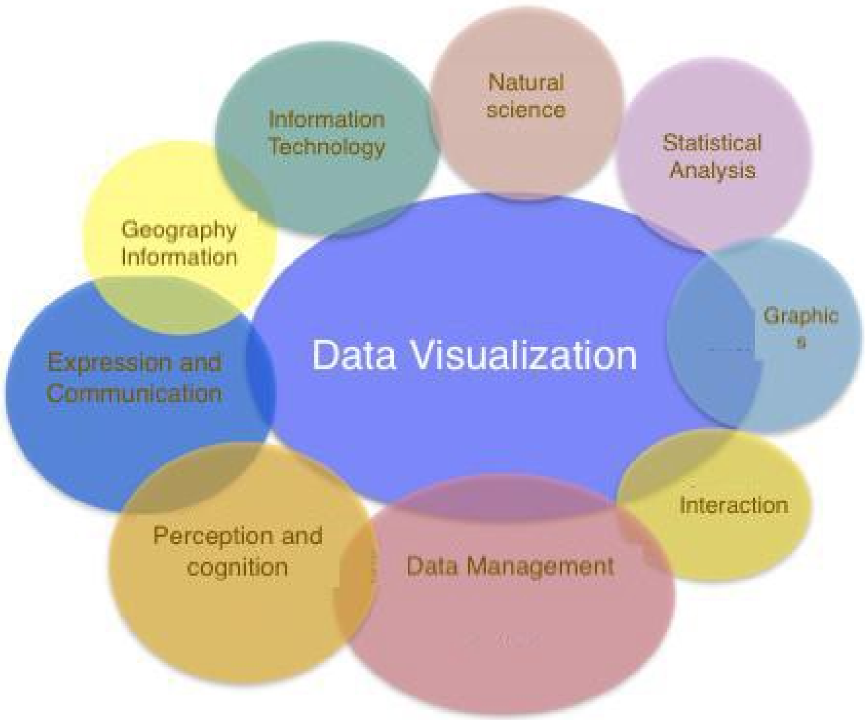


Figure 2.2: The field of data visualization

The following papers will present concepts, definitions, and tools in the field of data visualization.

2.4.1 What is Interaction for Data Visualization?

This paper tries to define what interaction means in the context of visualization, HCI, and how this definition can help further understanding and quality in the field of visualization [11]. They examine commonalities and differences between how interaction is viewed in visualization and in HCI and gain insights from several fields of visualization including information visualization, visual analytics, and scientific visualization with the input of researchers in the field of visualization.

The paper argues that the barrier to achieving the enrichment that is being called for in interactivity in visualization systems is not only a technical challenge but a challenge in the definition of interactivity for visualization [11]. To address this, they first define the current view of interaction, compare it to its view in the HCI community, then combine these definitions in order to broaden the scope of interaction in visualization.

In order to capture the current view of interaction in visualization the paper performs a critical review of papers in the field. Through this critical review, with the help of 22 visualization researchers, they ended up with 59 reviewed papers that attempted to define interaction in some way [11]. The definitions from these papers are as follows:

Mandatory components of interaction:

- **External entities**
  - The user
  - The data
  - Time
- **Internal entities**
  - Intent/Goal
- **External actions**
  - User-system dialogue
- **Internal actions**
  - Cognitive acts/Reasoning

External entities

The user is a human being who initiates the interaction, either end-users or designers. The data is an information source that is the user's main object of interest [11]. Other external entities can be physical objects such as mouse and keyboard, physical constraints, body or eye movements, speech, etc. The environment in which interaction occurs can also be considered an external entity, for example, whether it is a professional or casual setting, or whether or not there are multiple users.

Internal entities

Intent, or goal, can describe the level of data exploration or insight, need to acquire multiple perspectives on the data, etc. Intent can also mean how the user *intends* to use the data. Other internal entities are users' prior knowledge, skill, and ability when interacting with a visualization system. Interactions can also occur with an absence of intent.

External actions

'Dialogue', in this case, describes the exchange between the user and the visualization system. The user performs an action, and the system returns a reaction. Actions performed by the user can change the data or sometimes create new data.

Internal actions

Interacting with a visualization system involves cognitive acts from the user, performing reasoning or analytical processing on the data [11]. This means much of the interaction occurs internally in a user's mind.

The paper also notes reported benefits and critiques on interaction from these papers. Interaction was seen as necessary in order to handle increasing amounts of data, but that it has moved beyond its necessity [11].

> "It is now seen as a mean to amplify cognition in active, human-driven data exploration in which the user is in control of the information space. It is via interactive manipulation that 'knowledge is constructed, tested, refined and shared'." [11]

Some of the critiques concern how interaction is rarely the focus of research efforts in visualization, and when it is it is usually focusing on engineering or implementation rather than designing for interaction [11]. Other critiques highlight the limited focus on human and technology modalities, i.e. alternate ways of interaction other than keyboard and mouse. The lack of flexibility was part of the critiques, pointing to the constraints on user freedom in manipulating, inputting, organizing, and collaborating due to technological restraints.

In the HCI field, the paper presents these concepts that HCI uses to characterize interaction from their findings:¨

Dialogue

Similar to the definition in the visualization field, this concept describes the input and output between a user and an interactive system. Good interaction has direct, simple, and 'natural' dialogue with strong feelings of understanding and control [11].

Transmission

Transmission is the information being passed between the computer and the user. Good interaction maximizes the amount of error-free information being transferred.

Control

This concept focuses on minimizing error for a user target. Good interaction minimizes error, reduces the distance to user goals, and provides rapid and stable convergence to a target state [11].

Tool Use

Interaction is a set of tools that lets a user interact with a computer. Good interaction means useful tools that amplify user power.

Optimal Behavior

Optimal behavior refers to how humans usually optimize how they perform tasks, whether physical or mental. Good interaction allows users to maximize the use of their capabilities.

Embodiment

The concept of embodiment is the act of being and participating in the world, feeling that technology is a natural part of them. Good interaction uses artifacts in a way that assists the user in a non-disruptive way.

Experience

Experience is how users' expectations, reactions, and memories are a factor in interaction. Good interaction stimulates users' psychological needs in a satisfying way.

After comparing the difference in views in the visualization field with the ones in HCI, the paper makes the following compact definition of interaction for visualization:

> ”Interaction for visualization is the interplay between a person and a data interface involving a data-related intent, at least one action from the person and an interface reaction that is perceived as such.” [11]

The paper elaborates on this definition and discusses its limitations, but concludes that having a single concise definition will allow the field to create visualization systems that empower the users through interactivity.

2.4.2 Ten guidelines for effective data visualization in scientific publications

The paper aims to provide guidelines for visualizing data for scientific publications in order to promote effective conveying of information [28]. The main part of the paper consists of the ten guidelines for effective data visualization. Each guideline is explained in text and shown by a visualized example. The visualizations show bad and good examples of what the guideline is describing. The guidelines vary from general tips to advice for specific types of data. There is also given insight into common pitfalls in data visualization while describing these guidelines.

The ten guidelines are summarized as follows:

1. **Create the simplest graph that conveys the information you want to convey**

   This guideline warns of using too "flashy" graphs, as the main point of a graph is to present or explain information. They specifically mention the fact that visualization software often has functionality for making impressive graphs, and recommends curbing the urge to take advantage of these functions.

2. **Consider the type of encoding object and attribute used to create a plot**

   "Encoding objects" are described as the points, lines, and bars used in graphs. Which one to use and how they look (attributes) is important to consider based on what type of data is being visualized.

3. **Focus on visualizing patterns or on visualizing details, depending on the purpose of the plot**

   Presenting specific values is better visualized with bar or line graphs. When searching for patterns in data visualizations like heatmaps or bubble plots can be more useful.

4. **Select meaningful axis ranges**

   Limiting the visualization of the data to the relevant ranges improves the clarity of the graph. For example, if the data ranges from 2000 to 4000, starting the graph at 0 would create a lot of empty space and make the visualization less readable. This however is only relevant for specific types of graphs.

5. **Data transformations and carefully chosen graph aspect ratios can be used to emphasize rates of change for time-series data**

   Similar to guideline 4, using a logarithmic axis could help visualization, especially to show a rate of change over time. Altering the height/width aspect ratio of the graph could also benefit time-based visualizations.

6. **Plot overlapping points in a way that density differences become apparent in scatter plots**

   This guideline pertains specifically to scatter plot graphs. Scatter plots can be hard to understand if data is gathered together in densities. Decreasing the size or opacity of the data points can improve readability.

7. **Use lines when connecting sequential data in time-series plots**

   Plots that connect non-sequential data or values on either side of a period of missing data with a line imply a linear change between the points.

8. **Aggregate larger datasets in meaningful ways**

   Large quantities of data in a single visualization should be carefully presented. In this guideline, several techniques for doing this are presented.

9. **Keep axis ranges as similar as possible to compare variables**

When comparing different sets of data with different axis ranges the ranges should be displayed in a way that they don't hinder visual comparison. One solution could be to separate them into two graphs rather than having them in the same one.

10. **Select an appropriate color scheme based on the type of data**

    A color scheme can help support visualization. Color gradients varying from light to dark can emphasize low to high values. Contrasting colors can help to highlight opposing variables. Light, neutral colors and dark, stronger colors can differentiate between average to high values in a graph.

[28]

Since the aim of the guidelines is directed at scientific publications they can be useful for making visualizations for doctors and experts, but may not be as useful when considering patients and non-experts. While some of the guidelines are very technical and specific, most of them give advice that is useful and informative for visualization in general. In the conclusion, they acknowledge that the guidelines should not be treated as absolute rules, but that they will generally improve presentations of scientific data.

2.4.3 Big Data Visualization: Tools and Challenges

The paper addresses the increasing amount of data in the modern world, i.e. ”Big Data”, and challenges in analyzing, interpreting, and presenting large amounts of data [1]. Concepts include Big Data, Data visualization, Data analytics, and tools used in these contexts. Big Data is a name for the increased amount of data about most parts of our lives. Institutions like academics, governments, and hospitals are keeping increasingly detailed information about people. Most companies, IT-related or not, are storing all data they produce [1].

Ali et al. discuss why data visualization is important in the age of Big Data. They address some of the challenges related to it and they review some tools used for visualization. In the conclusion, they highlight the usefulness of visualization and how it helps keep track of large amounts of data in a more easily understandable way [1]. The tools reviewed are praised, but no ’winner’ is declared. They end the paper with suggestions for readers to consider their requirements and restrictions before choosing a tool.

The main takeaway from the paper is the tools reviewed. The different tools are presented clearly and their strengths and weaknesses are pointed out, making it a great resource for finding tools for data visualization.

2.4.4 User Interfaces for Search: Visualization in Search Interfaces

Reading and scanning text is a cognitively taxing activity when reviewing large amounts of textual information, and it must be done linearly [22]. Images, however, can be scanned quickly and information is perceived in parallel.

> ” A visual representation can communicate some kinds of information much more rapidly and effectively than any other method. Consider the difference between a written description of a person's face and a photograph of it, or the

difference between a table of numbers containing a correlation and a scatter plot showing the same information. " [22]

Visualization of information and data has become more and more commonplace in both media and business. This has lead to new creative innovation in the field, especially on the web. Several sites allow users to upload data and explore different ways of visualization. Visualizing abstract information can be difficult, especially textually represented information. Interactivity seems to be a useful property in visualizing abstract information. Functions like panning and zooming along with animated graphics help make information more digestible [22].

Web and mobile solutions can be used to facilitate visualization since all standardized browsers function on different types of operating systems and devices. This feature is of an advantage to any user making it possible to visualize and explore the data in a dynamic way.

2.4.5 Visualization Support for Clinical Care

There are several recent studies that speak in favor of visualization in clinical care by looking at its effects on the flow of care, decision support, and other activities such as data gathering, the difficulty of the data gathering process, cognitive load, time to task completion, errors, and improving situational awareness, compliance with evidence-based safety guidelines, usability, and navigation [29]. Good visualization is beneficial for improving quality and safety in following patient situations, compliance with treatments, clinical satisfaction, situation accuracy as experienced by nurses. It seems that collaborative effort and iterative development has been central to development of efficient dashboards with strong visualization [29].

A system called Medical Information Visualization Assistant v.2 (MIVA 2.0) was developed to support work in the intensive care unit. The testing of MIVA 2.0 was carried out using different types of questionnaires and semi-structured interviews [15]. Findings suggest that MIVA 2.0 has the potential to out-perform the use of paper charts in retrieving and analyzing patient data and has been appreciated for awareness of real-world intensive care unit activities. The evaluators notice its capacity to improve decision-making also via connecting it to the existing electronic patient record systems. The system seems to have secured a good balance between being informative and visually distractive [15].

There are more reports of how visualization helps decision-making in real-time. One positive effect is improved situational awareness that allows for rapid intervention in patient care and treatment [18].

All these results are of relevance for the visualization of this project since they are implemented within the real environment in which future users could consider making decisions based on efficient and customized visualizations.

2.4.6 Kaplan-Meier

Kaplan-Meier, also called Kaplan-Meier estimate, is a statistical method used for performing survival analysis for subjects after a given treatment [31]. The 'survival' times do not necessarily refer to an actual survival until a supposed 'death', but may refer to any event of interest where ’survival’ means the time-to-event [45]. While mainly used in medical research it is also used in other fields.

The data for a Kaplan-Meier survival analysis requires three variables for subjects being studied; the time, the status at a certain time, and the study group they are in [45]. If a subject drops out of the study, fails to show up to a follow-up, or their data is somehow invalidated or unavailable they are counted as ’Censored’ in the analysis instead of simply being removed from the results.

Results from such studies are presented in ’survival curves’, a visual representation of the points of data over a period of time. Figure 2.3 shows how these curves are presented. The horizontal lines show the survival duration, the vertical lines illustrate the change in survival probability at given events in the study [45]. These curves are usually accompanied by a table of the data, to allow readers to examine the data, as the curves usually aren’t detailed enough to allow reading data straight from a graph.

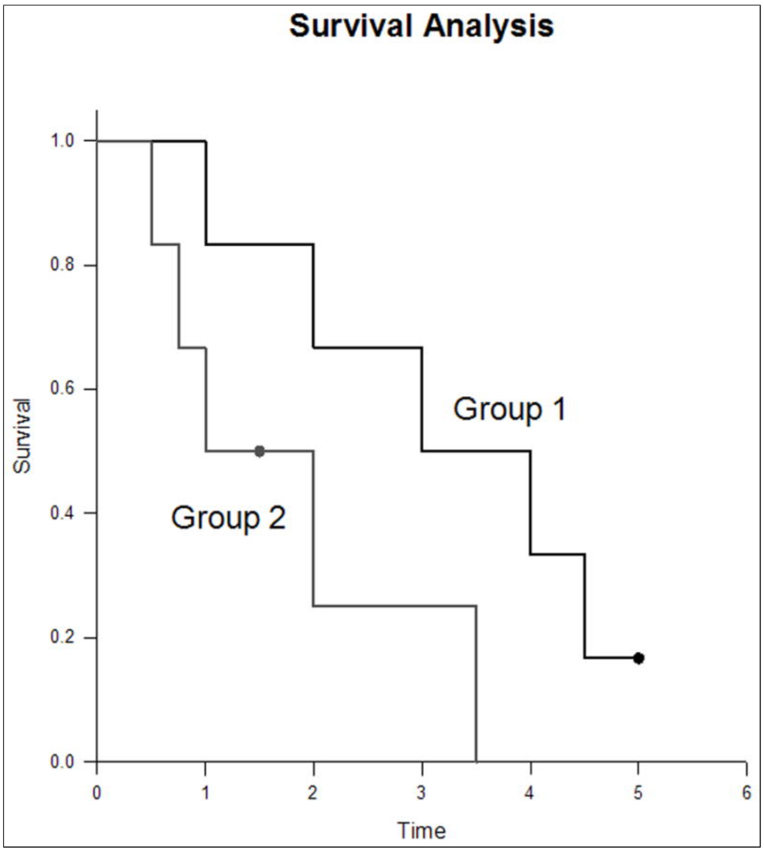


Figure 2.3: Example of Kaplan-Meier survival curve

The curves can be evaluated by different statistical tests that assess the results by comparing two or more curves with each other [31]. A ’log-rank’ test is the most common in Kaplan-Meier, and its results are usually presented alongside the graphs [45]. Another commonly used test is the ’Cox proportion hazard’ test, often referred to as just ’CoxP’. ’Proportional hazard’ refers to the risk of an event happening in a group compared to another group [31].

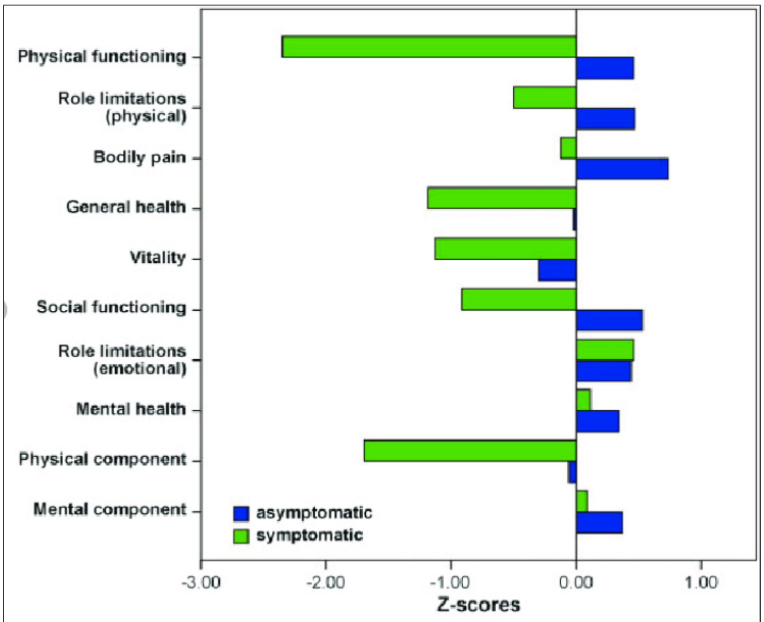
2.5 Patient quality of life

2.5.1 Health outcomes of patients undergoing cardiac surgery

This is a paper on a study that assessed health-related quality of life (HRQOL) for cardiac surgery patients before and after surgery [14].

The method they used was a questionnaire administered to patients at three different times; before surgery, at hospital discharge, and 6 months after discharge [14]. An Australian 650-bed hospital in Sydney was the hospital used for the study. Potential participants were given an information sheet, a consent form, and a pre-surgery questionnaire to take home before admission to the hospital for surgery.

The study used a questionnaire called ”SF-36”, a health questionnaire that takes 5-10 minutes to complete. According to Elliott et al. it is widely accepted as a reliable tool for measuring health statuses in populations [14]. Another questionnaire called ”15D” was also used. Also- a health-related questionnaire that takes into account the patient’s mental health, as well as general bodily functions and quality of life. Figure 2.4 shows the visualized results of an SF-36 study.



SF-36 scores. Graphical representation of the SF-36 quality of life scores for each of the eight domains and the two component scores (mental and physical component summary). Bars represent Z-scores that were calculated using reference values for the Dutch population, matched for age and gender.

Figure 2.4: Example visualization of SF-36 form data

The study sample was 101 patients over a period of 5 months [14]. Reported health statuses saw ”significant improvements” comparing the pre-surgery questionnaires to 6-months post-surgery. However, there was a negative difference when it came to mental health and social functioning. The paper ends by suggesting more should be done by hospitals post-surgery in helping patients recover, and further studies should be performed.

The paper gives insight into different health-related questionnaires and how they’re applied practically. The results about post-surgery mental health could be something to explore.

Some of the results are also visualized in graphs, showing how actual patient data is presented from a set of data.

2.5.2 Quality of life and functionality after total hip arthroplasty: a long-term follow-up study

In this paper, Mariconda et al. suggests more needs to be done in long-term follow-up for patients of hip-arthroplasty. They conduct a study of post-surgery arthroplasty patients using several different questionnaires and models to explore variables that affect patient health-related quality of life. They present two specific goals of the study:

- Evaluate whether patients who had hip arthroplasty more than 11 years ago suffered from severe functional impairment or disability.[36]
- Identify possible outcome predictors of long term quality of life and hip functionality post hip arthroplasty.[36]

The data collection tools used for the data studied in the study are questionnaires called ”SF-36”, Harris hip score, WOMAC score, and Functional comorbidity index [36]. They also used a questionnaire made specifically for the study. Using the data gathered from the questionnaires Mariconda et al. perform a statistical analysis of the data. They name several of the tests they used in this process.

> ”A two-sample t-test, ANOVA, and chi-square test were used to test the significance of the cross-sectional differences between groups. A Bonferroni test was used to test the differences between multiple groups. Pearson’s correlation coefficient was used to assess the relationships among patient-oriented outcomes.” [36]

In their results, they find that while there is some negative quality of life in the patients, it is still better than people who did not have hip arthroplasty at all [36]. They discuss some methodological weaknesses in their study in the conclusion, acknowledging a lack of a control group and the uncertainty of possible variations in the information-gathering.

Their methodology of analyzing medical data gives useful insight into how to approach a dataset of medical data. As well as several tools and models that are relevant to the dataset in this project. The discussion of methodological weaknesses brings up aspects that should be considered. The statistical analysis is thoroughly performed, but the presentation of results is basic. They are presented purely in tables of raw statistical data, making it difficult to understand without a thorough study of the tables. Graphical visualizations could have made the results more easily understandable, especially to non-experts.

2.5.3 Patient access to medical records and healthcare outcomes: a systematic review

In this article, the authors look at the effects of providing patients access to their own medical data and the effects of this[10]. They discuss the potential benefits as well as the risks, such as privacy concerns, improved understanding for doctors, benefits for patients, potential increased anxiety in patients, etc. While they conclude that there might not be enough benefits for patient access to become mainstream, they do provide a lot of information that could be relevant when deciding what/how to present medical data to patients.

The paper performs a review to determine what effect providing patients access to their medical data has on healthcare quality [10]. They review scientific studies from 1970 to 2013, filtering out sub-par studies using a quality of study rating form. The review found no negative impact on patient outcomes from access to health information, not even patient anxiety, contrary to a common expectation from physicians regarding the subject.

2.6 Related Work

This section presents theses that have worked on similar data, illustrating related topics and similar areas of research to this thesis. There are more theses that have worked on the NAR data, but these selected are ones that include forms of visualization of data, although none where these have been the focus of the research.

2.6.1 HALE, the Hip Arthroplasty Longevity Estimation system

In this thesis, Longberg [35] designed a system for total hip arthroplasty prosthesis longevity estimation. This system was developed to explore using machine learning techniques on biomedical data for expert users' (biomedical engineers and arthroplasty physicians) needs. As opposed to our thesis, the HALE system was designed only for these expert users, focusing more on technical research methods than how to present the data. The system interface is presented in Figure 2.5. While the system mainly returns data in tables and figures, there is one example of visualized data, as is shown in Figure 2.6.

Regardless of the limited scope of visualization, this system has created a user-friendly way to interact with the data. Through simple input, the user could search for the predictions of prosthesis longevity entering data that was relevant for the example at hand.

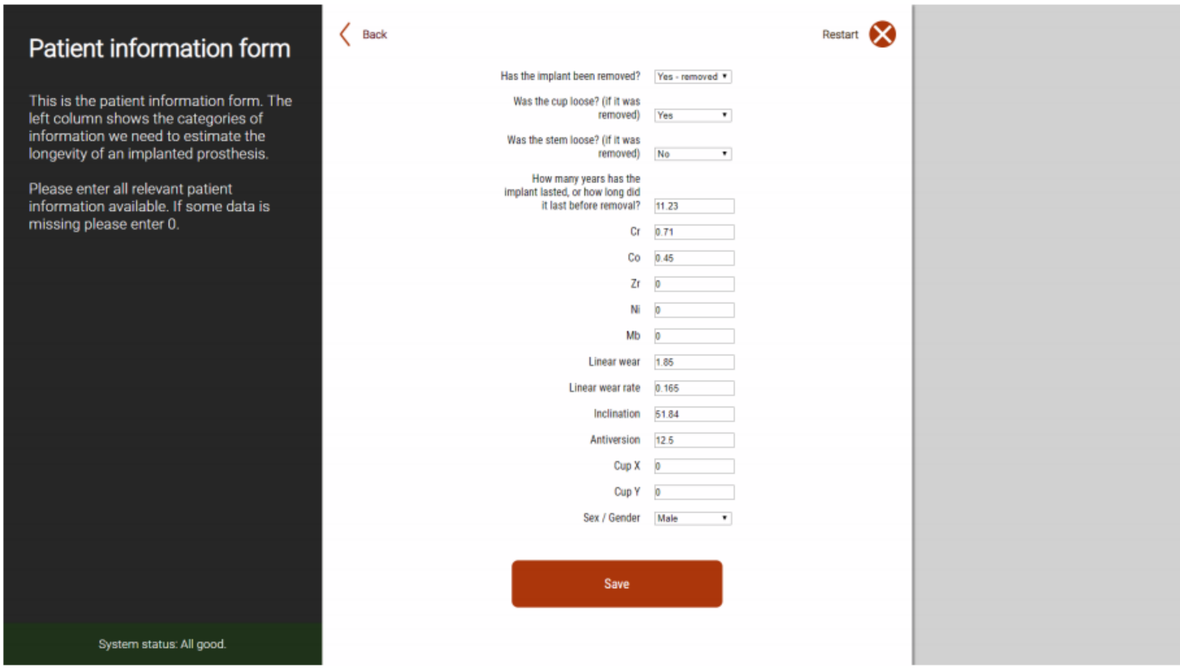


Figure 2.5: Example of the interface of the HALE system

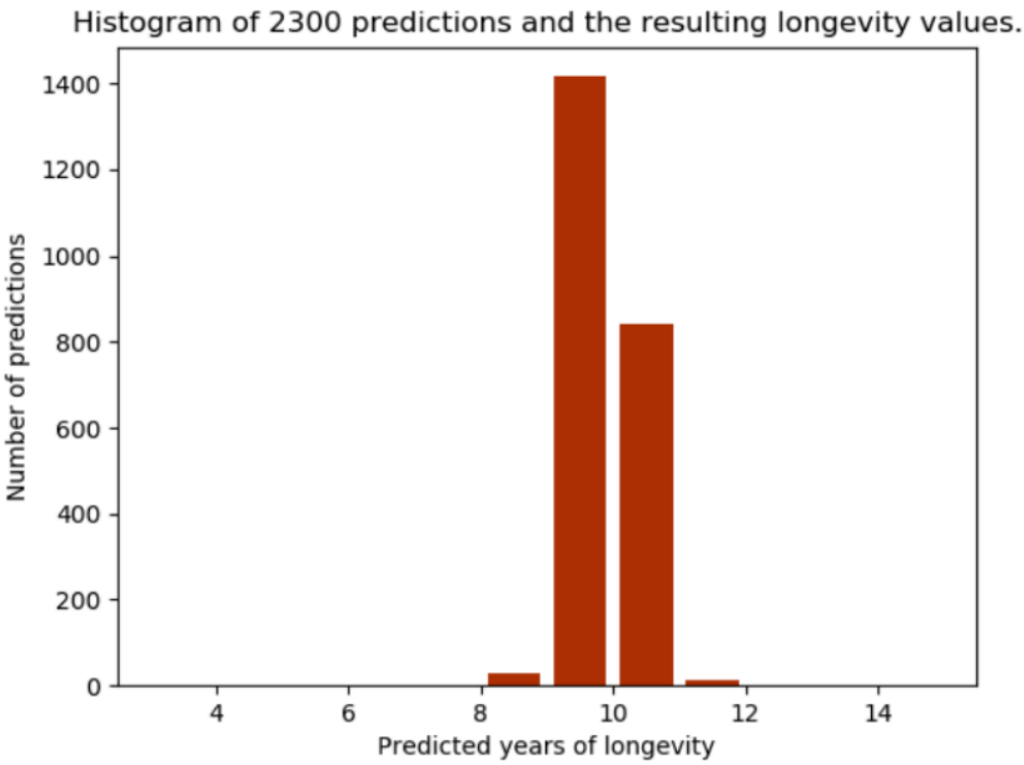


Figure 2.6: Visualization in the HALE system

2.6.2 Data Mining Approach to Modelling of Outcomes in Total Knee Arthroplasty

This thesis [25] seeks to provide insight into biomedical data beyond manual human capability by exploring the application of data mining methods on the knee arthroplasty data. The results of the thesis are a set of descriptive and predictive models for the data. While they did not create a system like Longberg [35], they present visualizations from the results of these models, as can be seen in Figures 2.7 and 2.8. These visualizations are of a very technical nature, but as the focus of the thesis was on the models and not the presentation of data to users, so this setup was understandable.

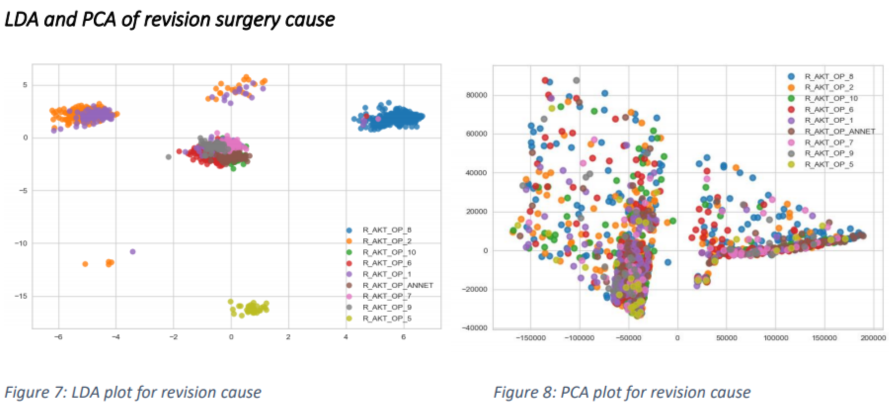


Figure 2.7: Iden: Visualized result of data mining method

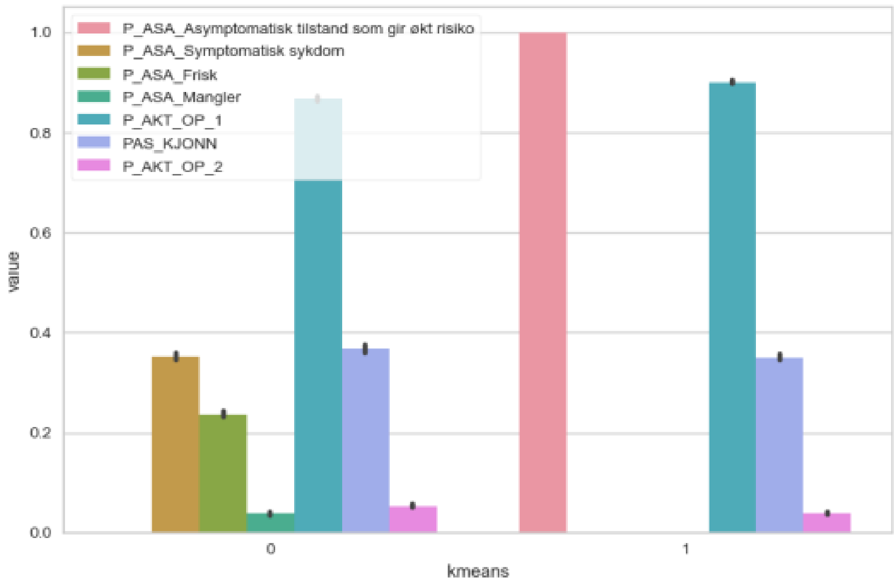


Figure 2.8: Iden: Visualized result of data mining method

2.6.3 Mining for individual patient outcome prediction in hip arthroplasty registry data

Similar to the previous thesis, the intention of this work [32] is to use data mining to develop models to predict patient outcomes from the register data. By using different data mining methods, the thesis explores the dataset and finds perspectives on the survival of different prosthetic devices in patient groups. Three learning algorithms were employed to examine possible prediction of data in the dataset, with varying degrees of reliability. Results from the research suggested that in order to perform more complex predictions the variables in the register data had insufficient explanatory power.

This thesis contains many visualizations, but they are mainly used to display an overview of the results of the dataset (Figure 2.9), and later to evaluate the results of the data mining methods rather than for sharing findings with users (Figure 2.10).

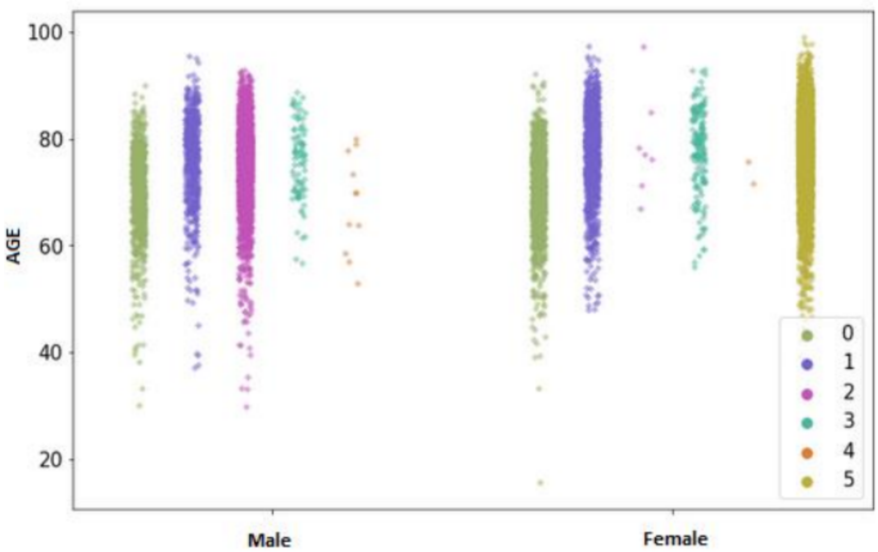


Figure 2.9: Kristoffersen: Visualization of dataset distribution

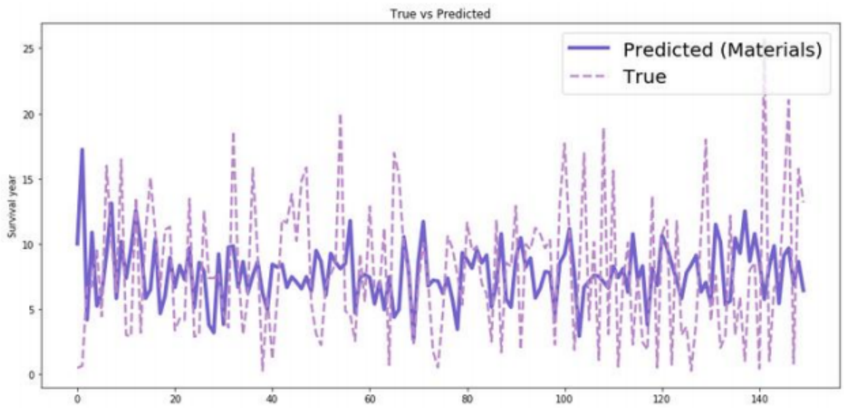


Figure 2.10: Kristoffersen: Visual inspection of model results

3 Methodology and Methods

This chapter will discuss the data and methods to be used in the project. The tools used in the thesis will also be presented and briefly discussed.

3.1 Data

The data used in this project are data samples from the Norwegian Arthroplasty Register (NAR). The register was started in 1987 and started collecting data about hip operations. In 1994 the register was expanded to include knee, shoulder, and other joint replacement operations. The Norwegian government has stated that hospitals must report operations to the register, with the condition that written consent from the patient is obtained [20]. The main objective of the register is stated to be the prevention of the use of poor prostheses in patients. The information recorded includes time and reason for the first operation, eventual revision surgery and reason for revision, type or brand of prostheses, patients' secondary illnesses, operation duration and location, and other technical operational details [20]. This information is collected through a standard form filled out by the surgeon.

The data is published in an annual report that combines the NAR with three other hip and joint-related registers. There are regular reports sent out to hospitals about their data and the national data, which is also published annually and made public every year at the register's website [20].

The register contains data for patient quality of life data. This data contains information about Health-Related Quality Of Life (HRQOL) post-surgery, i.e. their recovery, functioning, pain, social functioning, and mental health. While data mining in this dataset might not have as concrete pragmatic results as the arthroplasty dataset, it could provide interesting insight into patient recovery and the resulting impact on patient well-being. As suggested in [14], there is lacking research in this field. This data could be seen as a complement to the register data.

The visualization was running in parallel three other master theses projects that also utilized the same data sets. That meant collaboration in the team to define data mining and structure data mining processes on one hand, and the Human-Computer Interaction solutions on the other hand. As the dataset contained data from operations on several different joints, a preliminary division of labor was between hip and knee arthroplasty, with this project solely focusing on visualization.

The development of the different visualizations will be discussed further in Chapter 5.

3.2 Methods

3.2.1 Design science research

Design science research is a relatively new methodological approach for creating artifacts that serve human purposes and solve real-world problems, while also making a prescriptive scientific contribution [12]. This type of research creates artifacts that solve domain problems, or solution concepts, which must be evaluated by criteria of value or utility [12].

In the fundamental book [23], Hevner et al. present seven guidelines for design science in information systems research. The guidelines are based on the principle that knowledge and understanding of a design problem and its solution are acquired in the building and application of an artifact. The guidelines are:

1. **Design as an Artifact**

   "Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation."[23]. The creation of an artifact is the main result of design science research. These artifacts are however not usually meant to be used as a final product used in practice, but rather as innovations through which efficient and effective systems are designed and created.

2. **Problem Relevance**

   "The objective of design-science research is to develop technology-based solutions to important and relevant business problems."[23]. Since the objective of research in information systems is to gain knowledge and insights into the development and implementation of technological solutions, the goal of design science research is to create artifacts that solve problems and established challenges. Research must address the problems faced in the interaction of people, organizations, and information technology.

3. **Design Evaluation**

   "The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods."[23]. The environment in which the artifact is being developed needs to be the basis for how it is evaluated. The artifact can be evaluated by functionality, completeness, consistency, accuracy, performance, reliability, usability, organizational fit, and other relevant quality attributes. An artifact can be considered complete and effective once it meets the requirements and solves the problem it was meant to solve.

   Design evaluation methods:

   - **Observational**
     - Case study
     - Field study
   - **Analytical**
     - Static analysis

  - Architecture analysis
  - Optimization
  - Dynamic Analysis
- **Experimental**
  - Controlled experiment
- **Testing**
  - Functional testing
  - Structural testing
- **Descriptive**
  - Informed argument
  - Scenarios

4. **Research Contributions**

   "Effective design-science research must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies."[23]. Design science has three main potential types of research contributions based on the novelty, generality, and significance of the artifact. A research project must have one or more of these contributions. These contributions are:

   (a) **The Design Artifact**

   If it solves a previously unsolved problem, expands the knowledge base, applies existing knowledge in innovative ways, or produces a significant value the artifact itself can be seen as a contribution.

   (b) **Foundations**

   New constructs, models, methods, or instantiations developed through the research that extend or improve existing foundations in the knowledge base are important possible contributions.

   (c) **Methodologies**

   Creative development and use of the evaluation methods mentioned in Guideline 3 can also provide research contributions.

5. **Research Rigor**

   "Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact."[23]. While overemphasis on rigorous methodology could result in important parts of problems get abstracted away, it is possible and necessary for research paradigms to be both rigorous and relevant. Rigor is derived from effective use of the knowledge base and skilled selection of appropriate means of construction and evaluation. Evaluations have to be performed in

appropriate environments on appropriate subject groups. It is imperative to understand why an artifact works or not to enable new artifacts to be well constructed.

6. **Design as a Search Process**

   "The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment."[23]. In this context means are the resources available to construct a solution, ends are goals and constraints on the solution, and laws are uncontrollable forces in the environment. An iterative process that increases in scope through the iterations makes means, ends, and laws become more refined and valuable. Possible design solutions can be described as all possible means that satisfy all end conditions consistent with identified laws. It may not be feasible, however, to describe all possible relevant means, ends, and laws for a design. In these cases, a heuristic search strategy can be a solution to figure out *if* an artifact works well, although it may not show why it works.

7. **Communication of Research**

   "Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences"[23]. It is important that audiences are aware of the process in which the artifact was created and evaluated. It also needs to be understood by audiences with differing levels of knowledge, while still being advanced enough for experts. Presentation of design-science research needs to emphasize conveying the applications of the artifact to the relevant audience. Presenting these details in concise, well-organized appendices is an appropriate communication mechanism for a managerial audience.

3.2.2 Research Through Design

Research Through Design (RTD) is a research model for gaining contributions from interaction design other than new design methods [51]. The model attempts to complement existing methods in Human-Computer Interaction (HCI). One of the main focuses is on making artifacts that transform the world from the current state to a preferred state, in the model called the 'right thing'. Research with the model seeks to engage 'wicked' problems in HCI. Some examples of wicked problems are:

- *"The design of smart home services for families where parents address the paradox of wanting to care and protect their children while also wanting to make them independent and children face the paradox of desiring the comfort and security their home and family provide while also wanting to step out and discover and invent who they are and might be."*[51].
- *"The role of ubiquitous, assistive technology in aiding an elderly population to "age in place" in their own homes. It is wicked in that the stakeholders have conflicting goals including adult children who often want their parents out of the home in an environment that can better ensure their safety, and elder parents who have huge identity investments in their homes, and desire to remain, even when doing so creates tremendous social isolation."*[51].

The model integrates knowledge and theories from behavioral science, ’real knowledge’, with technical knowledge and opportunities from engineers, ’how knowledge’ [51]. Upfront research in ’real knowledge’ developed through a process of ideating, iterating, and critiquing potential solutions, leads design researchers towards the ’right thing’. This process produces concrete problem framing, artifact models, prototypes, products, and documentation of the design process.

Some of the contributions from this RTD model include:

- Identifying new opportunities for technology advancement.
- Providing research engineers with inspiration and motivation for what they might build.
- Helping identify important gaps in behavioral theory and models.
- Discoveries of both unanticipated effects and templates for bridging general aspects of the theory to more specific ones.
- Artifacts that provide concrete embodiments of theory and technical opportunities, which lead to practical applications of HCI research.
- Holistic research contributions that reveal framing of problems and balances between intersecting and conflicting perspectives.

With the focus on the production of artifacts, the artifacts created in this model provide subject matter for discourse and continuing of conversation in the HCI community [51]. Created research artifacts can be more traditionally evaluated in order to search for similar approaches to common problems, making the artifacts potential pre-patterns for new design patterns.

Design researchers following this model departs from traditional design researchers in several ways [51]. They work more similarly to design practitioners, addressing under-constrained problems. It is not meant to replace the other design research roles in HCI, but rather add one that allows design researchers to work more as a collaborative equal with other HCI researchers.

The model distinguishes between research artifacts and design practice artifacts in two important ways [51]. The research intends to produce knowledge for the research and practice communities, not a commercially viable product. This focuses research by reducing effort made in considerations of economics in manufacturability and distribution. The focus remains on creating the ’right thing’. The second distinction is that research contributions should be artifacts that demonstrate significant invention. They should be novel integrations of theory, technology, user need, and context instead of merely refinements of existing products. There must be significant advancement demonstrated through the integration. Meteoric technological advances in hardware and software drive the invention of novel products in HCI, more so than any other interaction design domain.

RTD also has some criteria for the evaluation of interaction design research. This will be elaborated upon in Section 3.3.

3.2.3 System development method

The design of the artifact employed Agile development. By designing, testing, and evaluating prototypes through an iterative process that increased in fidelity over time the design secures a solid foundation. The specific method used is called 'Kanban', a 'lean' agile workflow management method [27]. Kanban is most often used by the use of a 'board' that keeps track of the progress of the project. The board separates work into three main sections; work to be done, work being done, and work that has been done. An example of such a board can be seen in Figure 3.1. Some of the benefits of Kanban are that it can be applied to an ongoing project without disrupting workflow, it encourages continuous small incremental changes and discourages large sweeping changes, thus keeping the process agile, and the visual of the board makes it easy for a team to keep track of the workflow [27]. Kanban was used to plan and execute the work which was shared and overlooked by the whole team of four master student developers.

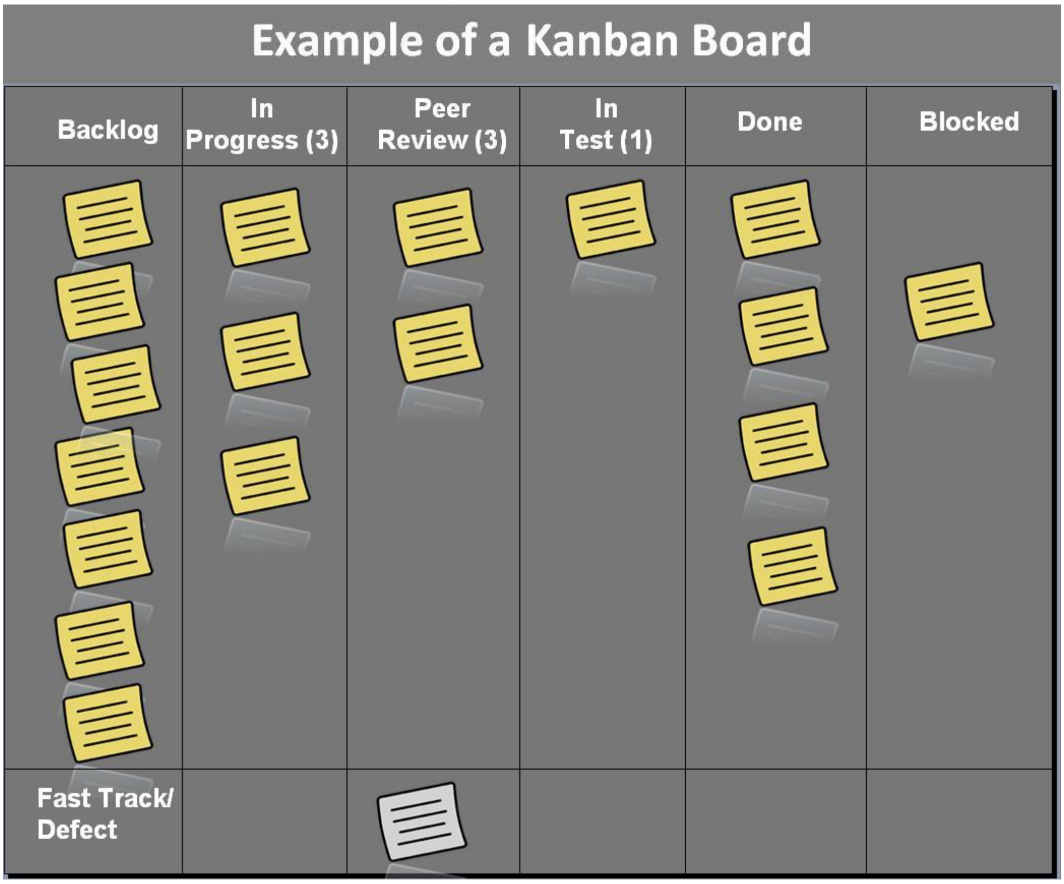


Figure 3.1: Kanban board example

3.2.4 Data mining

The importance of preparing and cleaning data prior to running data mining operations was highlighted in [50]. The article also asserts that it takes up a majority of the total data processing as data cleaning remains instrumental in data mining. Data mining procedures capable of analyzing data are numerous and readily available in software packages such as scikit-learn [43]. Two recent master theses have used several multivariate statistical methods to analyze outcome and longevity of the knee and hip prostheses [32, 25]. For example, they used cluster analysis, correlation analysis, etc. This thesis has used additional methods such as Kaplan-Meier survival analysis and utilized visualization to present data from a demographic perspective (map of Norwegian counties) as well as an extended presentation of demographic data.

3.2.5 User friendliness

Visualization could be a means to user-friendliness by enabling a good experience viewing and interacting with the data. There are several well-established guidelines to follow to secure a good quality of user-friendliness and design [46, 39]:

- The ease of learning how to use a system
- How to operate it
- How to prepare the input data
- How to interpret the results
- How to recover from errors
- Accessibility
- Help features
- Ease of use

In this thesis, an effort was made on preparing and operating the data, as well as on ease of use. Assuming that the solution is robust and the user group is familiar with the clinical questions no dedicated help features were developed, and the same goes for the interpretation of results.

3.2.6 Interactivity

There are many definitions of interactivity. Four categories of interactivity are proposed: the Data, the Data Representation, the Temporal Dimension, and Contextualizing Interaction [9]. There is also a notion of the powerfulness of interactivity which is namely low (zooming, data orientation), medium (navigation, toggling), and high (filtering, highlighting, linking, etc.).

There is restricted interactivity that includes viewing and browsing activities such as looking at the data, zooming, and scrolling. Then there is a more complex level of interactivity which is afforded by a database interaction. This includes interacting with the content and different layers of information. In this thesis, we tried to achieve more complex interactivity using the register data.

3.3 Evaluation methods

The following subsections present methods used to evaluate the different aspects of the project.

3.3.1 Usability testing

Usability testing is often used in iterative development as it provides an overview of aspects that work or not, and a controlled way of measuring improvement between iterations [33]. In usability testing, users are asked to perform tasks in an interface in a controlled setting. This eliminates single-user issues while highlighting problems, the potential for improvement, and insights into how users interact with the interface. In addition to testing users, usability testing can incorporate an 'expert review'. In an expert review experienced UX experts examine the interface in order to eliminate common/basic problems before testing on non-expert users.

There are some criteria for evaluating interaction design research within HCI [51]. Zimmerman et al.'s RTD model proposes four criteria for evaluation of contributions in this field: process, invention, relevance, and extensibility.

- **Process**

  There is no guarantee that repeating the same process will produce the same result. The rigor applied to the methods and reasons for selected methods are the more important parts of the work. Documentation of the contributions needs to have detailed information about this when describing the process.

- **Invention**

  Contributions must demonstrate that they have produced novel integrations that address specific situations. A literature review that details the aspects that demonstrate how the contributions advance the state of the art needs to be performed. Interaction designers also need to show how advances in technology could result in significant advancement when discussing integration as an invention. This is what communicates guidance to the HCI research community on what to build.

- **Relevance**

  Scientific research has a focus on validity. Scientific work must be documented well enough so that peers can reproduce the results. In design research, however, the same problem will often produce different artifacts. The focus then shifts from validity to relevance. Interaction design researchers need to frame the work within the real world and articulate the preferred state the design attempts to achieve and why it is preferred. Neglecting to do this negates the real-world impact of the research.

- **Extensibility**

  In this context extensibility is the ability to build on the outcomes of interaction design research. Well-described and documented design research lets it impact future design problems and increases knowledge created by artifacts.

Questionnaires and surveys

Questionnaires and surveys use metrics such as the Likert scale for questions. The Likert scale asks users where they fall on a scale, often 1 to 5, 7, or 9 [33]. There are free online tools, e.g. Google Forms [26], that make it easy to create and share questionnaires online. However, these questionnaires can be hard to spread to a significant amount of users and are susceptible to non-serious answers.

3.3.2 Data mining

The data mining process benefits from evaluation as well. Osborne [42] discusses several tests that can be done to evaluate the quality of the distribution of the data before data mining begins:

- **The Ocular Test**

  This test simply involves looking at a graphic representation of the distribution of the data. It is the most common and useful method researchers use. It should never be the only method used, bust should always be the first, as humans are good at understanding nuances visually than by simple numbers.

- **Examining Skew and Kurtosis Figures**

  Skew refers to the symmetrical nature of a variable. If one end of the distribution of data is disproportionately higher, the data is said to be *skewed*. *Kurtosis* is a similar aspect of describing differences in the 'shape' of a graph distribution. Skew and kurtosis have standard variables in statistics. Comparing the dataset to the standards is a good way to evaluate the distribution.

- **Examining P-P Plots**

  P-P plots are examinations of the actual data in comparison with the expected theoretical normal distribution. Deviations from these comparisons can indicate errors in the data. Similar to the Ocular Test, however, P-P plots should only serve as a visual inspection tool. There is no clear rule for how much deviation is 'too much'.

- **Inferential Tests of Normality**

  These kinds of tests examine whether a variable conforms to a predetermined type of distribution, or to what extent it does not. Two tests mentioned of this type are Kolmogorov-Smirnov(K-S) and Shapiro-Wilk(S-W).

Osborne [42] mentions that these tools may not always point to the same conclusion, and it is up to the discretion of the statistician to utilize all information at hand to make the best decision.

3.3.3 SUS

The System Usability Scale (SUS) is a form for measuring usability. It is a 10 item questionnaire with a 'Likert scale' 5 point scale from completely disagree to completely agree. The ten questions are:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

When calculating the SUS score, there are several ways of measuring it, it can be grades, adjectives, and percentage, etc [34]. The ten questions are structured in a way that makes the odd numbers positively loaded questions and the even number negatively loaded. When calculating the score, the odd numbers subtract 1 from their value and the even numbers subtract their value from the number 5. After calculating the final scores, the next step is to multiply them by 2.5 to get the SUS score out of 100. It is important to remember that the result is not a percentage, but a SUS score. The research has established that a SUS score of more than 68 is considered a good score [34].

3.3.4 Nielsen's Heuristics

Nielsen's heuristics are 10 usability heuristics for creating good UI. The method's goal is to find usability problems in the UI during design iterations. Heuristic evaluation utilizes a small set of evaluators to examine the UI with the usability principles [41].

These are the 10 principles or heuristics one should follow during evaluation [40]:

1. Visibility of system status. The system should always provide users with information about what is going on, through appropriate feedback within a reasonable time frame.
2. Match between system and the real world. The system should speak the language of the user, rather than using system-oriented terms. Information should appear in a natural and logical order.
3. User control and freedom. The system should support undo and redo functions since users often choose system functions by mistake.
4. Consistency and standards. Users should not have to question the meaning of words, situations, or actions that mean the same thing.
5. Error prevention. Either good error messages or diligent design that prevent problems from occurring.
6. Recognition rather than recall. Lessen the user's memory load by making objects, actions, and options visible.

7. Flexibility and efficiency of use. Accelerators, which the novice user will not see, may speed up the interaction for an expert user, therefore the system can cater to both inexperienced and experienced users.
8. Aesthetic and minimalist design. Rarely needed or irrelevant information should not be present in dialogues.
9. Help users recognize, diagnose, and recover from errors. Error messages should be expressed in plain language and no codes, giving the user a firm indication of what the problem is and suggest a solution.
10. Help and documentation. Using a system without the need for documentation is often the “best” way. However, sometimes it may be necessary to provide help and documentation focusing on the users’ tasks.

3.3.5 Content Evaluation Table

An evaluation questionnaire was developed to evaluate the content of the back- and front-end systems combined. Since the users are expected to work with the whole system, the content was seen from the users’ perspective. In parts, it addressed the data mining methods and how to present their results, and also it assesses components of the visualization as well as the user interface. The questionnaire was based on ten components to make it as concise as possible. It also offered a possibility to offer comments and suggestions.

A Likert scale starting from ’Totally Disagree’ to ’Totally Agree’ was used to quantify the value for each component.

These ten component are as follows:

1. Choice of Data Mining (DM) tasks
2. Need to add additional tasks
3. Welcoming starting page with something like demographics
4. Save all DM sessions
5. Choice of visualization
6. Level of interactivity in visualization
7. HCI outlay is satisfying
8. Need to add additional HCI functionality
9. HCI interface is well suited for experts
10. HCI interface has potential to meet patient needs

This thesis will focus on the components relevant to the visualization aspect of the questionnaire.

3.4 Tools and Technologies

This section takes a look at the different tools used during the project. They were selected as they were the most suitable for development in this thesis. A brief description is given to illustrate how they were applied.

3.4.1 JavaScript

JavaScript is a programming language, primarily used for the web [38]. It is the main tool used for visualizations in this thesis. It was selected because it didn't require learning a new language as it was familiar to us. It is also lightweight in both programming, testing, and presenting results. JavaScript is often used in data research and has countless community-created 'libraries' that can be implemented for use with a single line of code by simply linking to an external source, and are often completely free to use. There are also more professional libraries created by companies or organizations where payment is often required in order to utilize them, at least for commercial use.

CanvasJS

CanvasJS is a JavaScript library for creating charts and graphs for HTML5 [16]. It offers a range of different types of graphs and charts with various optional dynamic and interactive functionality. It was chosen for this project for its ease of use, comprehensive interactive functionality, selection of charts, flexibility, and because it is free to use for students.

3.4.2 Python

Python is a high-level general-purpose programming language [17]. Python was initially meant to be the main tool utilized in this project is due to its extensive collection of scientific libraries [43]. Due to the collaboration with the back-end students [24, 52] where they performed the data mining, this project's reliance on Python became less relevant. It also turned out that while Python's libraries are extensive, the libraries in JavaScript had better functionality for dynamic interactivity, where Python's were better at static visualizations. Python was, however, useful for manipulating data received from the back-end students [24, 52] in preparation for visualizing. And as their theses used Python as one of the main tools, it was also used when accessing data through the API they created.

3.4.3 Trello

Trello is a free, web-based tool for visually managing projects in Kanban-style boards [2]. It allows users to create task boards with different columns and freely move tasks between them. Being web-based, it easily allows for several users to remotely access and utilize boards in a collaborative setting.

3.4.4 Inkscape

Inkscape is a free, open-source vector graphics editor [44]. It was used in this project to edit SVG files.

3.4.5 IDE

For HTML, CSS, and JavaScript technically only a text editor is required, however, there are several tools for these languages that make development more efficient. An IDE (Integrated Development Environment) helps keep the code organized and has functionality for debugging and compiling. The IDE used in this thesis was Visual Studio Code [8].

4 Requirements

This Chapter presents ethical considerations and appropriate approval that was obtained from the Norwegian Centre for Research Data. The target groups are physicians, researchers, and patients for whom requirements were gathered by understanding current practices in the arthroplasty register data. Some data was received from patients in form of answered disease-related questionnaires (Section 2.4.5).

4.1 Ethical considerations

This research has been approved by the Norwegian Centre for Research Data (Norsk senter for forskningsdata - NSD). All participants involved have given their signed informed consent. The approval from NSD is in Appendix A. Interview guides and an informed consent form and can be found in Appendix B. All the research participants were made aware that they can remove themselves from the research at any point in time and that their privacy would be secured. No sensitive personal information would be requested.

4.2 Target Group

All major target groups were represented by personas. These personas are:

- Physicians and Orthopedic surgeons
- Medical Researchers
- Arthroplasty patients

From the visualization point of view, we have assumed that physicians, orthopedic surgeons, and medical researchers have a deeper knowledge of the data and procedures regardless of their particular work experience. We focused on providing expert users data that will document the back-end data mining procedures and showcase their most interesting results. Patients should have insight into the outcomes of surgery and their reported quality of life. They are not expected to gain detailed knowledge of the data, but rather an understanding of their own situation in terms of quality of life and some basic information about their prostheses (knee and hip).

The user groups are described in more detail as personas in the related front-end masters thesis [49].

4.3 Research Participants

4.3.1 Users

Users have been recruited at the arthroplasty register in Bergen based on the previous research projects and consisted of surgeons, biomedical researchers, and statisticians. An initial meeting with the register provided insight into standards of treatment, current practices of publishing annual reports, and statistical analyses. The biomedical research laboratory has provided information about their particular focus on revision surgeries and detailed analysis of explanted prostheses. The meeting helped illustrate what kind of visualizations of the results are common both in the literature and in the annual reports.

4.3.2 Usability Experts

Five usability experts participated in the evaluation of the artifact. The majority of these had a degree of Master of Information Science, with four to ten years of professional experience.

4.4 Establishing Requirements

There are two main sets of requirements: functional requirements that secure a functioning artifact that can demonstrate the feasibility of the future product, and non-functional that provide additional features that could make the artifact more efficient but are not operationally necessary.

4.4.1 Functional

It was necessary to understand what potential needs users could have and what data coming out of the analysis could substantially benefit from the visualization. Since the input is data coming from the back end development the following requirements were functional:

- Ability to import data from the back end developers' API
- Ability to produce understandable data visualization
- Secure and define metadata to explain visualizations
- Export visualizations at user request
- Export data to external software systems with additional HCI features
- Provide additional detail in form of frequency, standard deviation, and distinguishing patient groups
- Interactive user exploration whenever applicable

4.4.2 Non-functional

Non-functional requirements considered the aesthetics of the visualization aspects.

The solutions were:

- User friendliness
- Aesthetically pleasing to look at
- Designed to fit the needs of the front end development (HCI interface) [49]

5 Visualization Development

This chapter will present the development of the visualizations for the NAR data samples. It gives details of the procedures and methods used to develop the visualizations and starts by reflecting on the front-end collaboration.

5.1 Group work

As mentioned previously, this project has been cooperating with three other masters' theses. The common denominator was the data from the Norwegian Arthroplasty Register and a goal to develop a potential system working with the register. The system would enable performing data analysis through a user-friendly interface with the help of visualization. We were divided into "back-end" and "front-end" sides of development; the two data mining theses being the back-end, and this project and the HCI thesis being the front-end.

The group cooperation was performed through the agile system development method Kanban with the use of the web-based tool Trello to keep track of the development progress. Kanban has two main elements: visualizing workflow and limit work-in-progress. As can be seen in Figure 5.1, Trello helps do precisely this by dividing the work visually into different 'cards' that separate the work that has been done, the work that is being done, and the work that is going to be done.

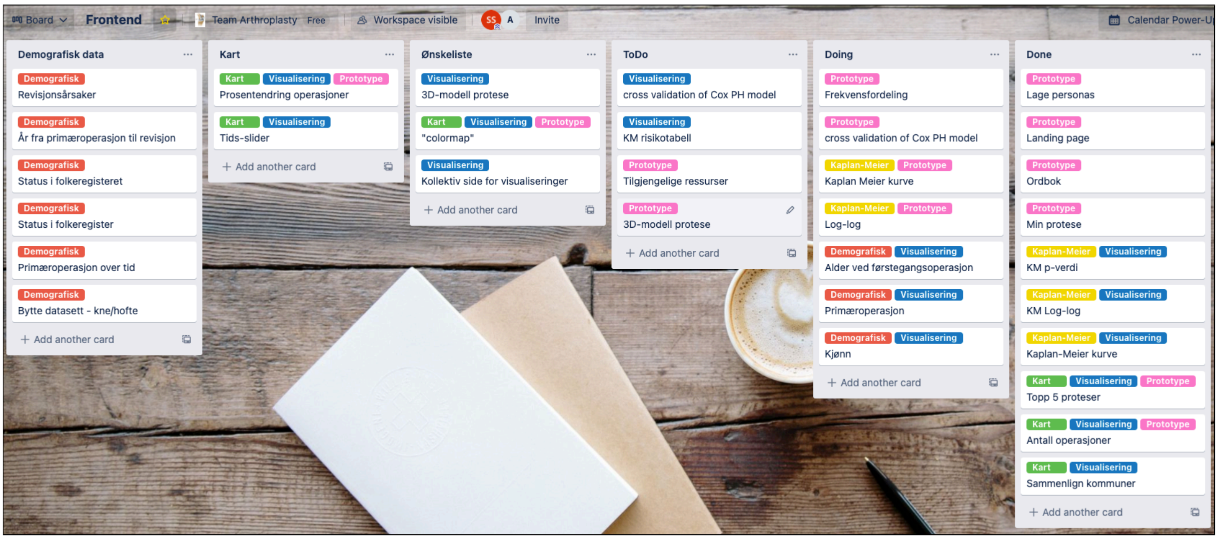


Figure 5.1: Front-end Trello

5.2 Iteration 1

The first iteration of development started with the idea of visualizing data using a map of Norway. The idea came from the fact that much of the data samples contained data related to different counties in Norway, so we thought that this type of visualization could be of interest to users across the user groups. Early in the project the tool D3 (Data-Driven Documents) [4] was chosen as the main tool for visualizing. Once we started experimenting with making an interactive visualization of a map of Norway, D3 turned out to be a much more cumbersome tool than anticipated.

After some research into map visualization was performed, an alternative solution was found which did not use any JavaScript library, but instead used vanilla JavaScript functionality in order to manipulate an SVG file of a map of Norway. This solution was found through a Google search that led to a piece of code posted on CodePen.io [7], a site for sharing code online. The author of this code had marked it as free to use to anyone (Appendix D), and the code was used as the basis for what became the final map visualization. The original code was done to visualize a map of a different country, however, the source of the map also had maps of Norway available. The source of the maps was SimpleMaps [48], which offer their maps for free (with some caveats, see Appendix D). Figure 5.2 shows the very start of the map visualization once the code had been downloaded and adjusted to the Norwegian application.

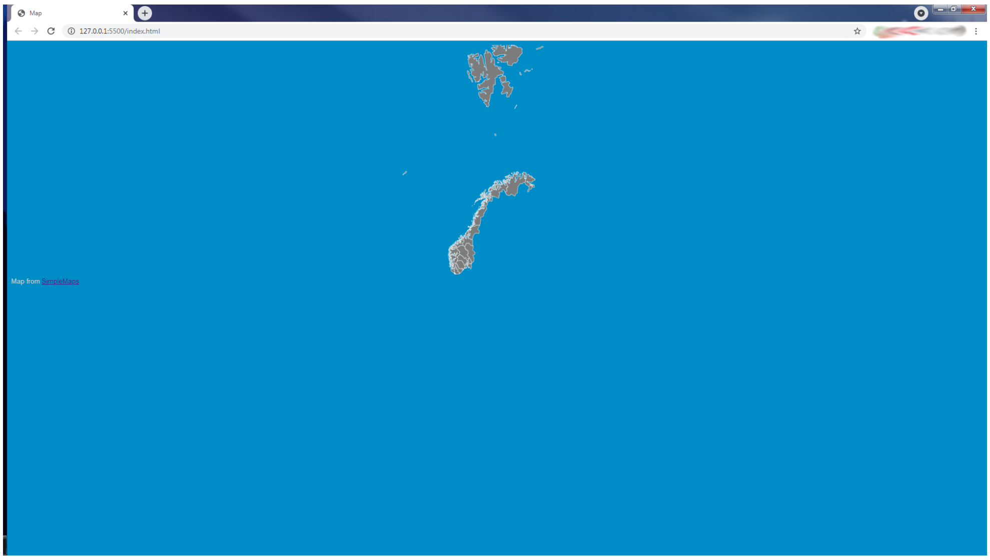


Figure 5.2: Start of map visualization

At this point, the map had the basic functionality of hovering over counties to see their names (since the SVG contained that data), but no other data was presented as the data from the code was meant for a different country.

5.3 Iteration 2

Once the tool and environment were in place, work started on enhancing functionality and entering data for the map. The first task was adding relevant data to each county. The master students working on the data mining [24, 52] were relatively early in the process at this stage and did not have all the relevant data available yet, but they had performed some processing and one of the available datasets was the change in amounts of operations performed in the first half of the sample-data and the last half (chronologically). This data was the first to be visualized on the map. Once this data was added the focus shifted to functionality and aesthetics. The code already had interactive functionality like panning and zooming, color change on hover, and a 'popper' that displays relevant data on hover. This functionality only allowed a user to examine one county at a time, however. This, along with the fact that researchers were one of the user groups, inspired a feature allowing a comparison between two different counties. It was implemented by having two 'info boxes' next to the map displaying data from counties as selected by the user. The functionality of this feature, as well as other interactivity on the map, will be presented in further detail in Chapter 6.

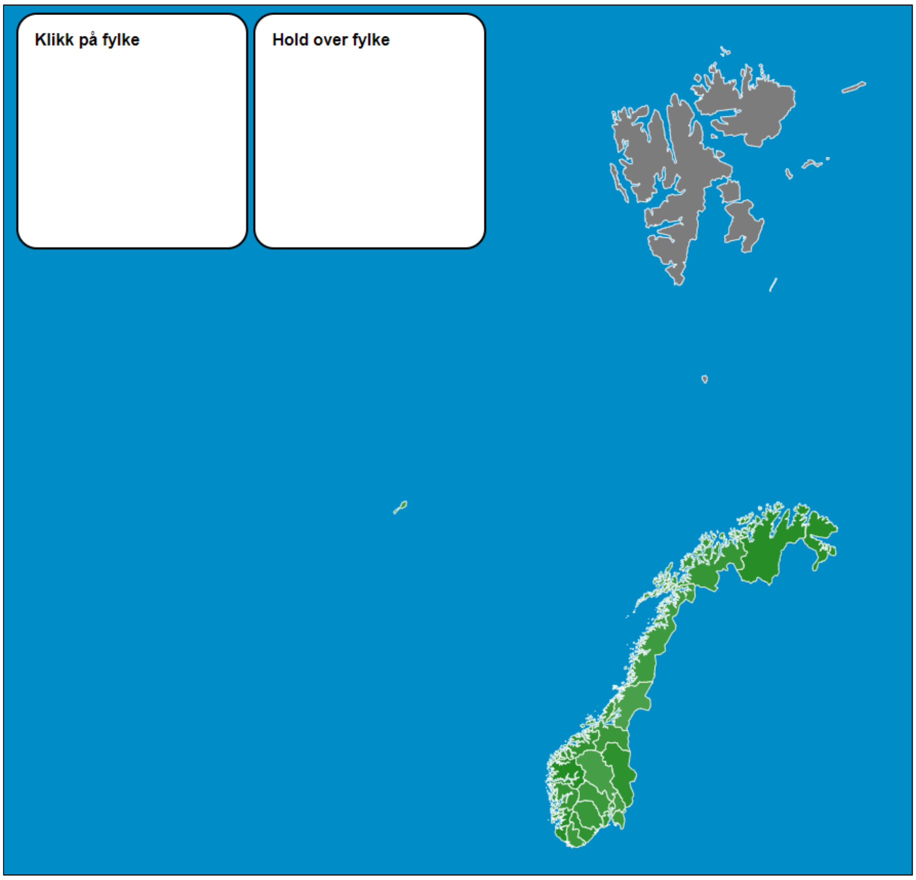


Figure 5.3: Early version of the map

Figure 5.3 shows how the visualization looked after the infoboxes had first been implemented. As it can be seen in this Figure, the territory off the northern coast, the archipelago Svalbard, is grey. This is because the data sample does not contain data on Svalbard. Because of this and since it was affecting the scaling of the map, it was decided that it should be removed from the map in the visualization. This was done by editing the map SVG file with the free, open-source vector graphic software Inkscape [44] which is compatible with SVG files. During this process, it was discovered that the map also included an uninhabited Norwegian island in the South Atlantic Ocean called 'Bouvetøya' (Bouvet Island). It turned out this island was the main culprit of the scaling issues, so it was removed along with Svalbard during this process. Figure 5.4 shows the map after implementing these changes and having added some data for the counties.

Figure 5.4: Map after editing

During this iteration, the back-end students [24, 52] had progressed and now had data ready for Kaplan-Meier from the data samples. Kaplan-Meier is a feature in the annual reports from the NAR, but they are static graphs representations, making them a natural candidate for interactive visualization. As D3 had already been rejected due to its complexity, research started on other visualization tools. Several JavaScript libraries for visualization were found, including one that had existing functionality for Kaplan-Meier graphing, but most of them were not free to use. CanvasJS was chosen because it was free to use and due to its wide range of different graphs with extensive documentation and relevant guides on its use. As CanvasJS does not have built-in functionality for graphing a Kaplan-Meier survival curve we needed to utilize the existing graphs in the library. The most suitable one turned out to be a 'Step Line' chart, which is a line chart where the lines between data points are connected with exclusively horizontal and vertical lines instead of direct or curved lines. Since this is how Kaplan-Meier curves are presented it turned out to be a satisfactory solution. With data on different materials used in hip prostheses, the first version of the Kaplan-Meier visualization was created. Figure 5.5 shows this stage.

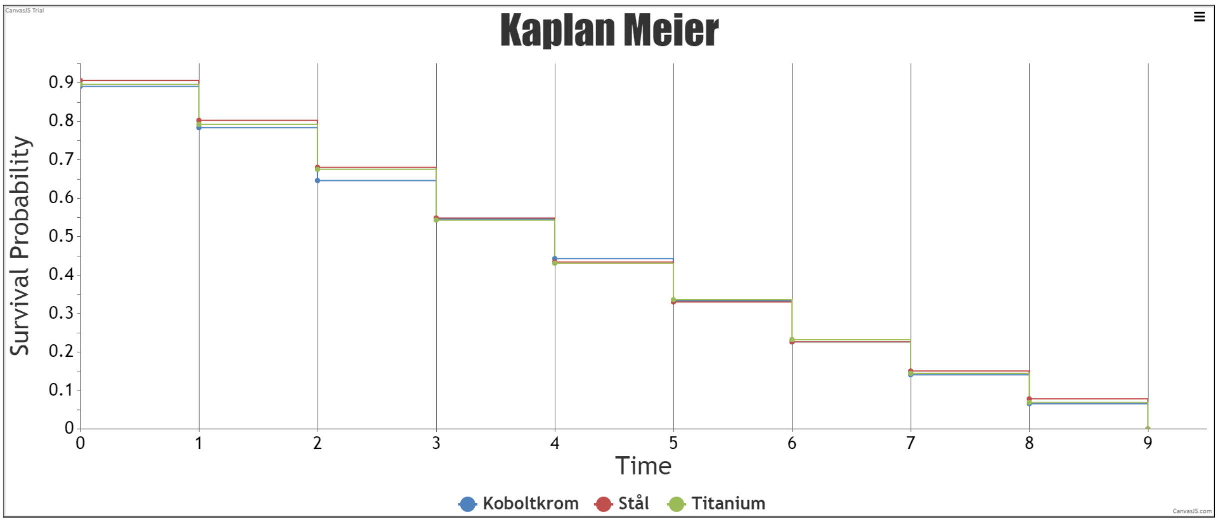


Figure 5.5: First version of Kaplan-Meier visualization

5.4 Iteration 3

The county map and Kaplan-Meier visualizations had the basic functionality in place, so work started on implementing further functionality and adding general improvements. Work started first on Kaplan-Meier as it was the last one worked on in the previous iteration and was most in need of improvement. An important part of displaying Kaplan-Meier survival curves is showing the 'confidence interval', i.e. the estimated accuracy of the lines in the graph. CanvasJS allows for the combination of different graphs in one chart, which allowed us to approximate the confidence interval in our chart. A 'range area' graph was added to the chart overlaying the existing lines. While the confidence interval is also supposed to use horizontal and vertical lines, but this was somehow not working at this stage, as it can be seen in Figure 5.6. It turned out to be a bug, but a workaround was implemented, which can be seen in the final version in Chapter 6.

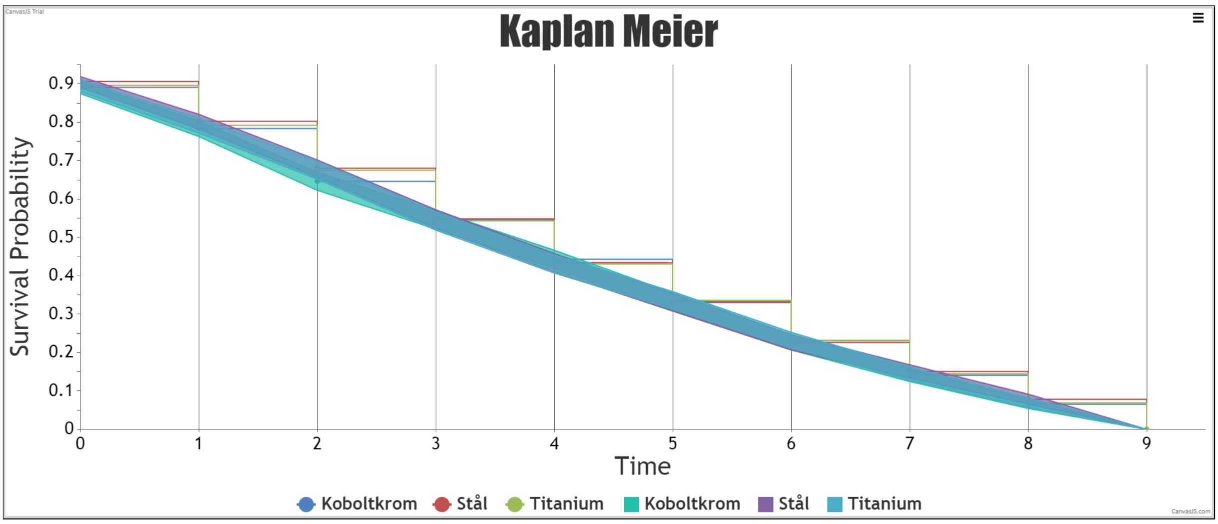


Figure 5.6: Kaplan-Meier curve with survival estimate

In this iteration, the student doing the HCI interface [49] became involved in the design of the map. Through this cooperation, several new types of functionality were implemented. A slider was added that allows users to scan across a time period for relevant data. Zooming was restricted to not allow the user to zoom too far out and lose the map, which was an existing problem. Elements were relocated on the page and the page as a whole got a general redesign. Figure 5.7 shows how the county map page looked after this cooperation.

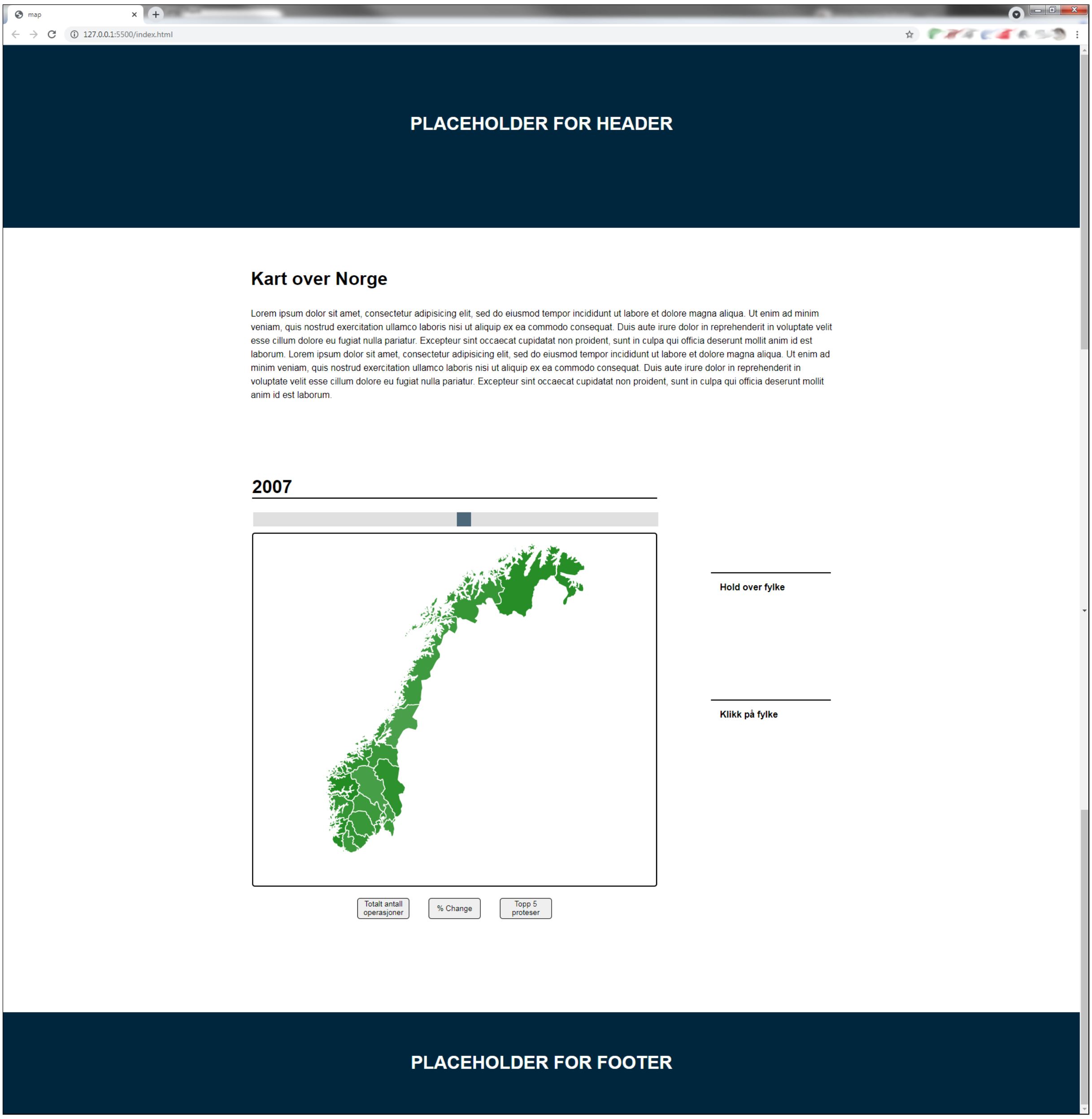


Figure 5.7: Version of map after redesign

As the interface developer [49] moved on from making the HCI prototype in HTML/CSS to an Adobe XD version, this became the only direct cooperation on the visualizations.

5.5 Iteration 4

As the design and functionality of the map were at a satisfactory level, we started working on retrieving data directly from the API delivered by the back-end developers [24, 52]. Up until this point the data had been used in the visualizations with JSON files of data loaded with the code itself. By connecting to the API we were able to import the data directly from the source through JavaScript 'get request' that exported the requested data. These functions were implemented specifically for this purpose by the back-end developers for use-cases like this. This connection and data retrieval from the API was proven functional and would be useful for a practical system. It was not used in further iterations, however, as having to start and keep the API running was more cumbersome than simply using the JSON files directly. It was also functionally identical in this small-scale development. This cumbersomeness was naturally a result of the API being in development alongside this project and not an inherent flaw of the system.

The next visualization was a collection of several different types of demographic data about patients. As we were getting more familiar with the data at this point more angles were becoming apparent. The idea came from the example of the county map in which the data could be of interest among the user groups. Creating these visualizations was a much more efficient process as we were already familiar with both the data and the tools. CanvasJS was used once more as it had the selection of different types of graphs, something that was suited for visualizing a collection of different data. Different graphs like bar, line, and pie were used to visualize data about patient age, gender, number of operations, and census status. The visualization at this stage can be seen in Figure 5.8.

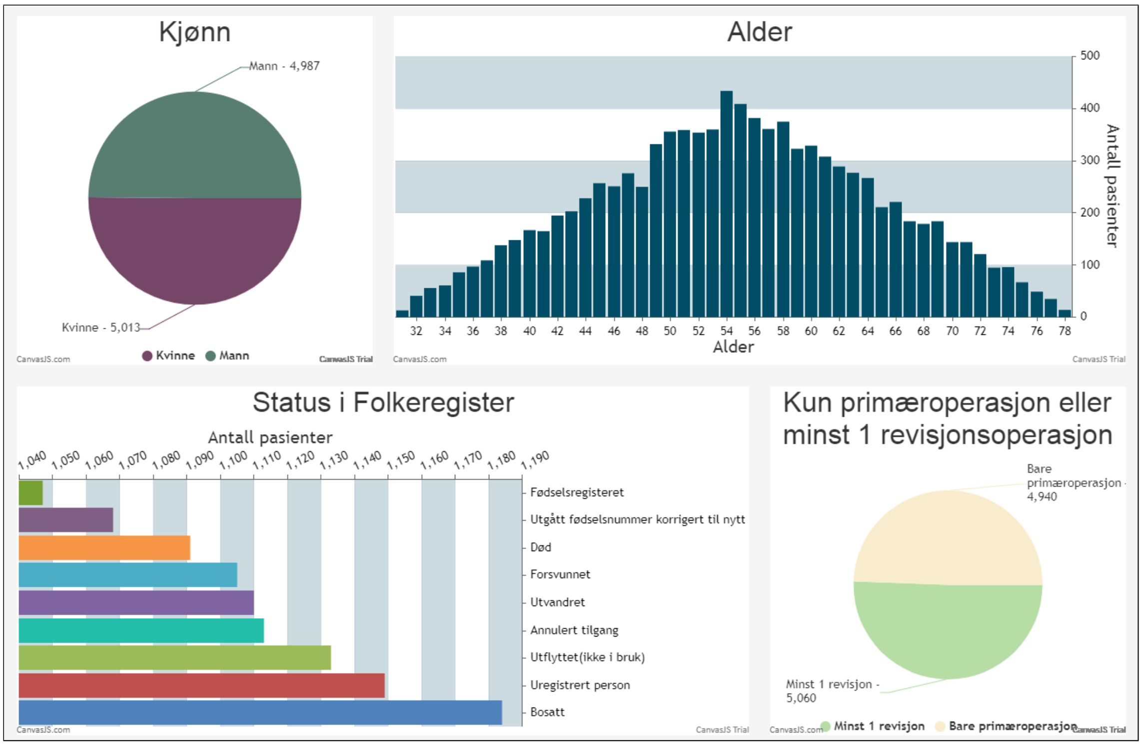


Figure 5.8: Early version of the demographic data page

With the framework in place, it became fairly trivial to add new visualizations to the demographic page. As we had gained at this stage access to data about both hip and knee patients, a function was added that switched the data being presented between these two, while keeping the same graphs. Figure 5.9 shows the full page with a few more graphs added, and a header with two buttons for switching data.

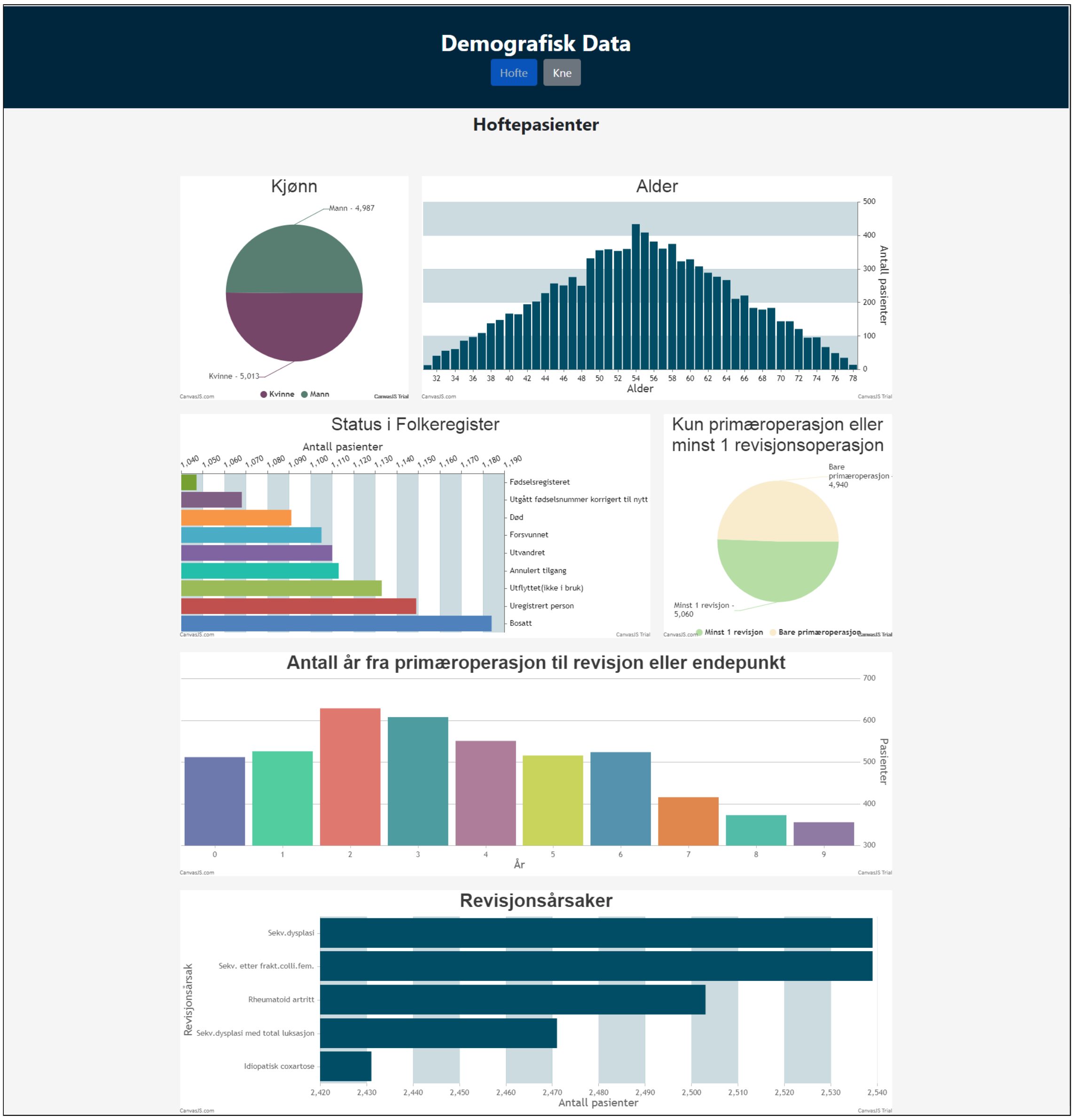


Figure 5.9: Demographic data page for hip arthroplasty patients

6 Results

This chapter presents the final visualizations in more detail, describing their functionality and design with accompanying images displaying them. As the visualizations were made for the Norwegian register data, all the text in them is in Norwegian.

6.1 Visualizations

6.1.1 County Map

The data from the register contains information about where patients have undergone operations, which is listed as ”Fylke”. This inspired the visualization of a county map of Norway displaying the information for each county as this information could be of interest to several user groups.

Functionality

Figure 6.1 shows the basic look of the map when loading the page. At the top is a slider with the chosen year displayed, defaulting to 2007. The box around the map keeps it in place. On the right are two info-boxes that display information when interacting with the map. They are initially blank, but the information is presented once the user selects a county.

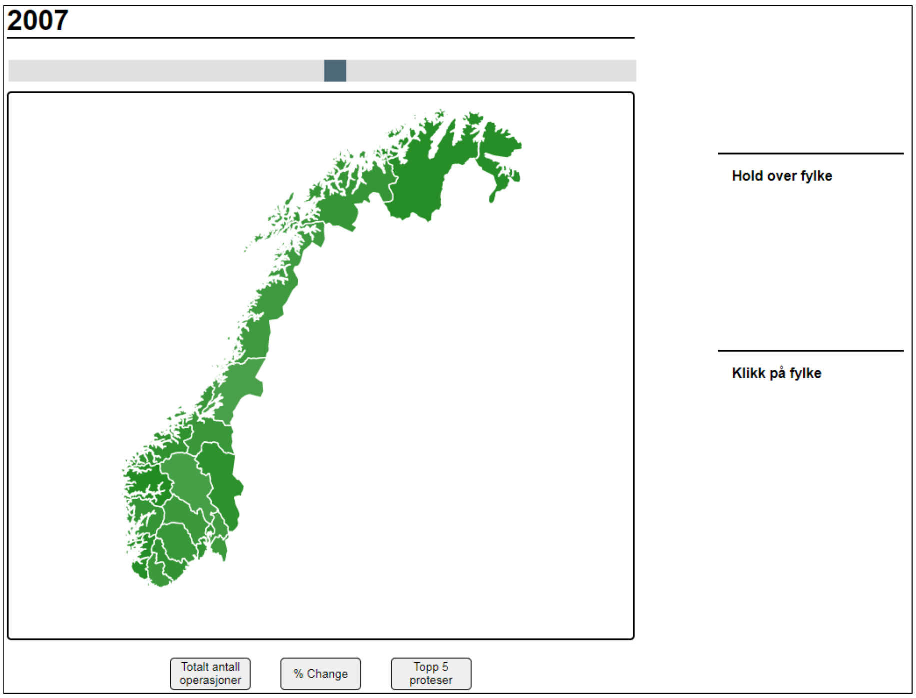


Figure 6.1: Map Startstate

At the bottom of the page are three buttons that select different types of data presented on the map. For example, one of the buttons allows a comparison of the top five prostheses in two counties.

Figures 6.2 and 6.3 show the difference in the slider position. In this case, it's showing the number of operations performed per county in the chosen year. The color of each county changes based on the data, going from a lighter to a darker color to denote higher numbers.

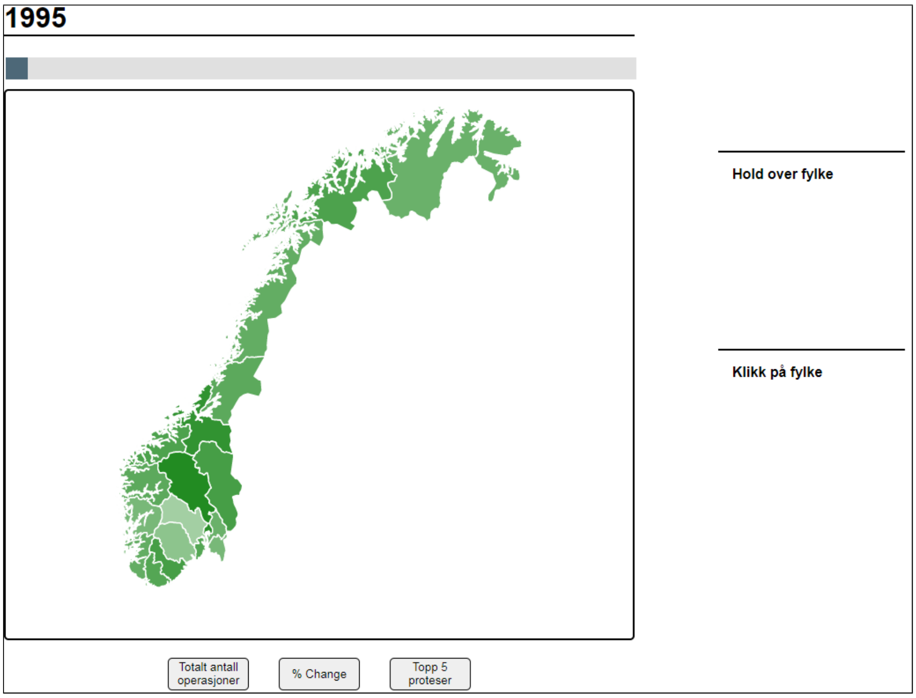


Figure 6.2: Map Slider at 1995

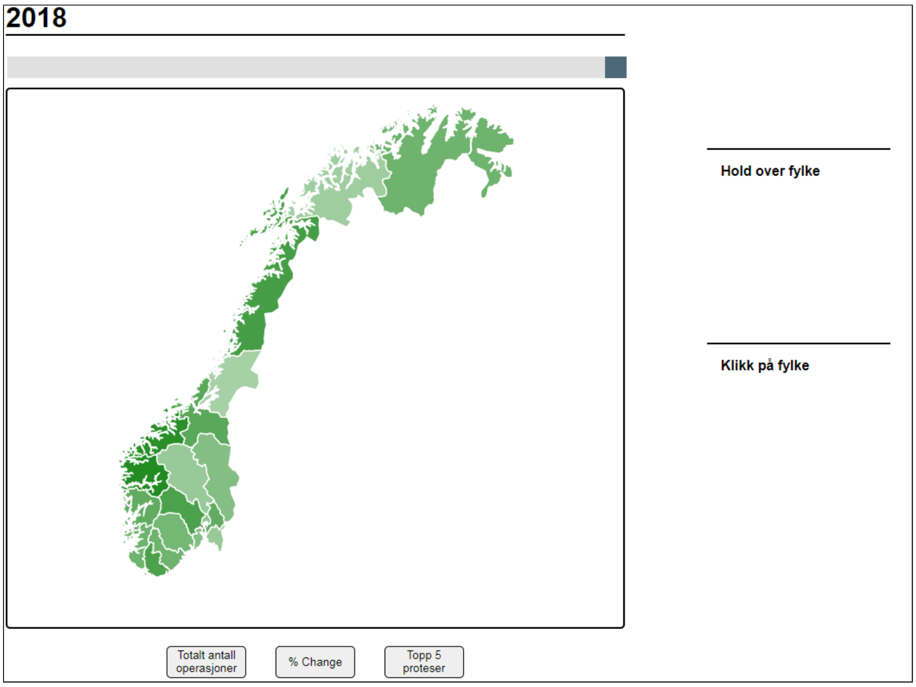


Figure 6.3: Map Slider at 2018

In Figure 6.4 one of the counties has been clicked and the cursor is hovering over another. When hovering over a county it changes to a blue color and the chosen data is displayed in a 'popper' next to the cursor as well as in the top box on the right. When a county is clicked, it changes to a darker blue color and displays the relevant data in the bottom right box. This allows the user to compare the data of two different counties. In this figure, the data displayed is the number of operations performed in 2007 registered in the register.

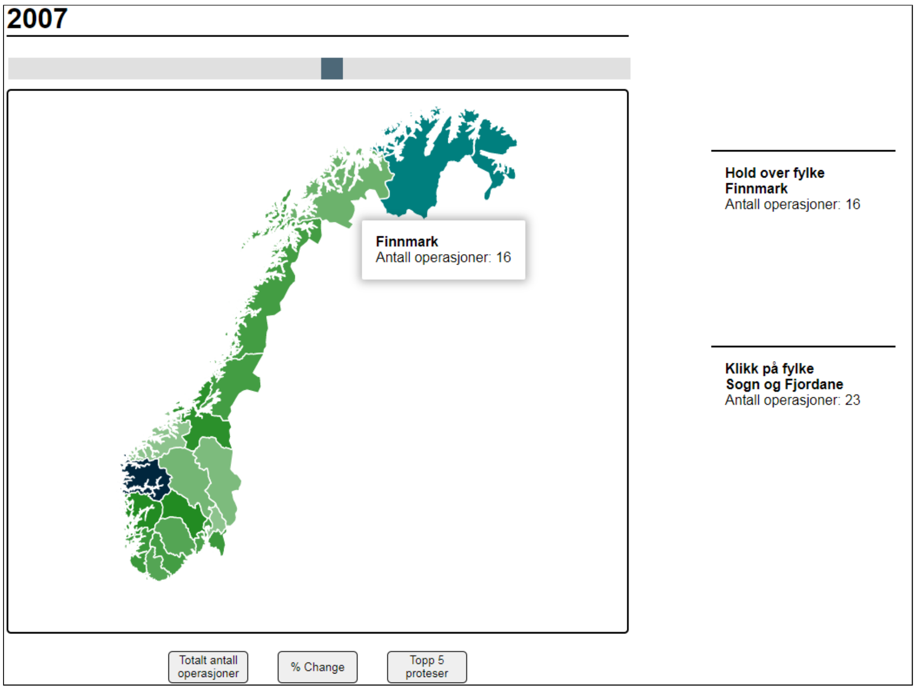


Figure 6.4: Comparing counties

Scrolling when hovering over the map allows the user to zoom in and out on the map, making it easier to navigate to the smaller counties on the map. When zoomed in it is also possible to pan around the map by clicking and dragging. This can be seen in Figure 6.5.

Clicking one of the buttons underneath the map changes what data is displayed when interacting with the map. Figure 6.6 shows the rightmost button is pressed and the change in data highlighted.

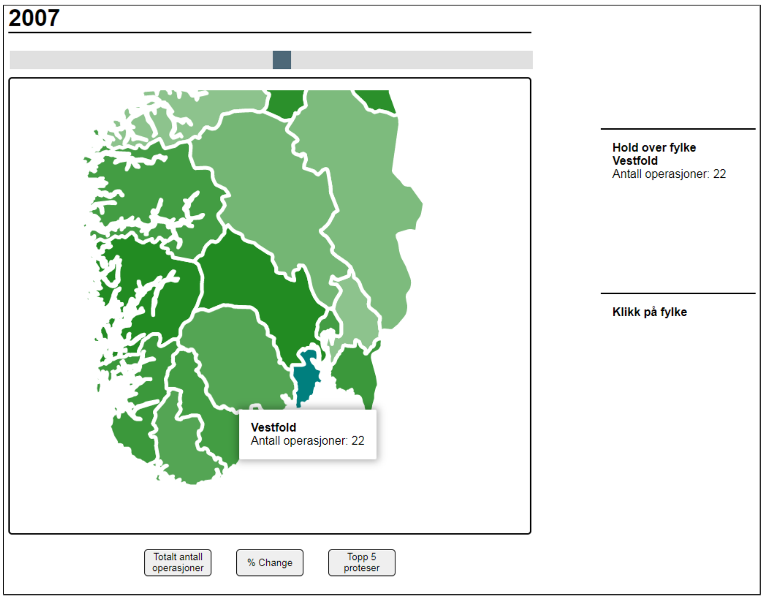


Figure 6.5: Map zoomed in

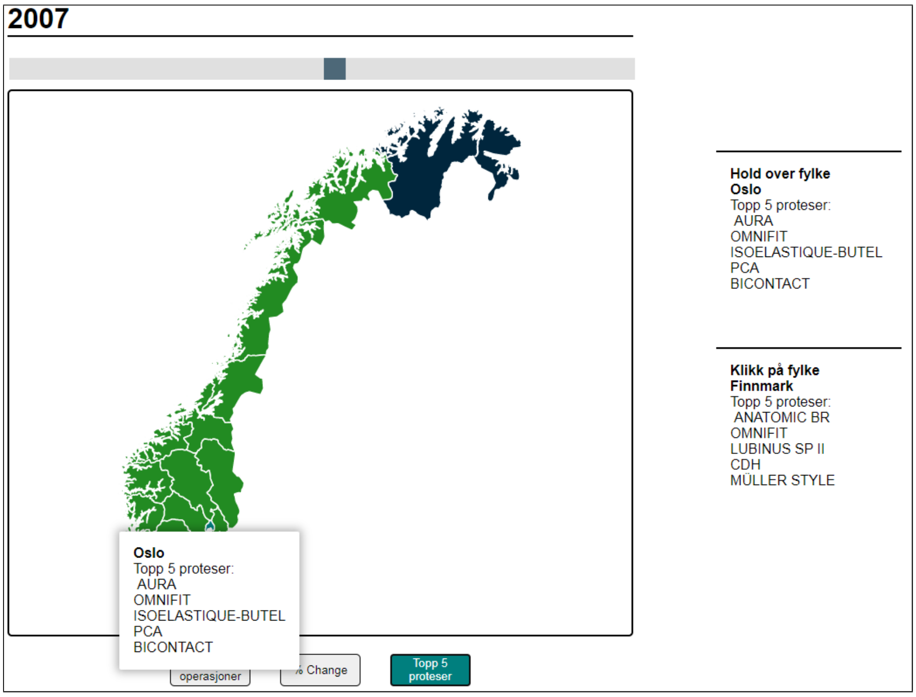


Figure 6.6: Selecting and viewing top five prosthesis data

6.1.2 Demographic data

The data samples from the register contained several types of demographical data for the patients. This inspired the following visualization showing this data in different types of graphs.

Figure 6.7 shows the page of the demographic visualizations as presented when the page loads. In order to see the rest of the visualizations the user will have to scroll down the page. Figure 6.7 shows data about Hip Arthroplasty data, while Figure 6.8 shows Knee Arthroplasty patients.

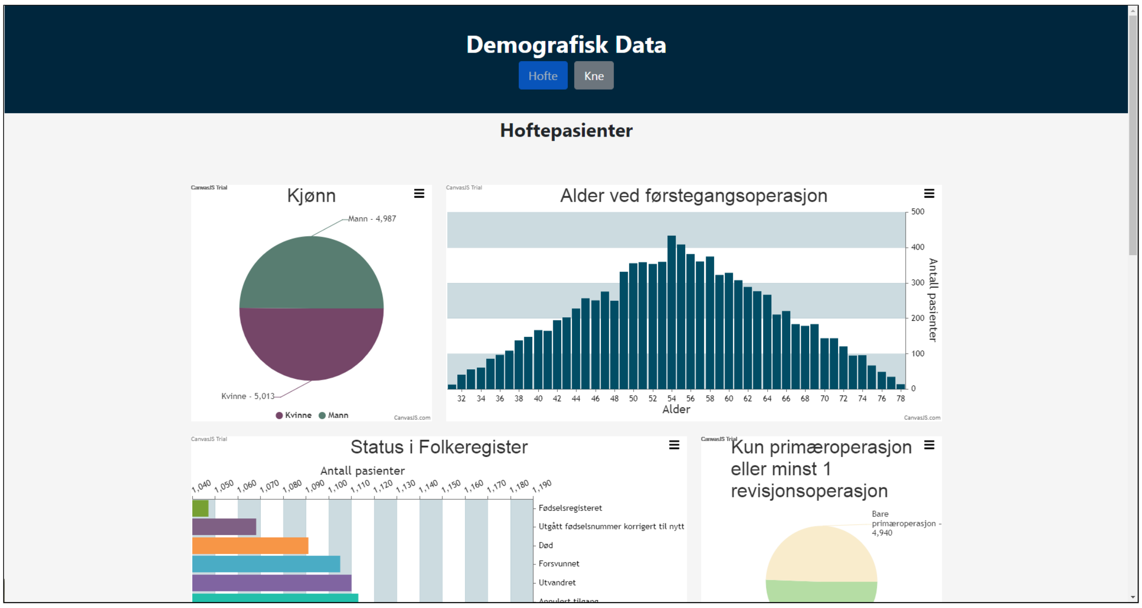


Figure 6.7: Demographic data page for hip arthroplasty patients

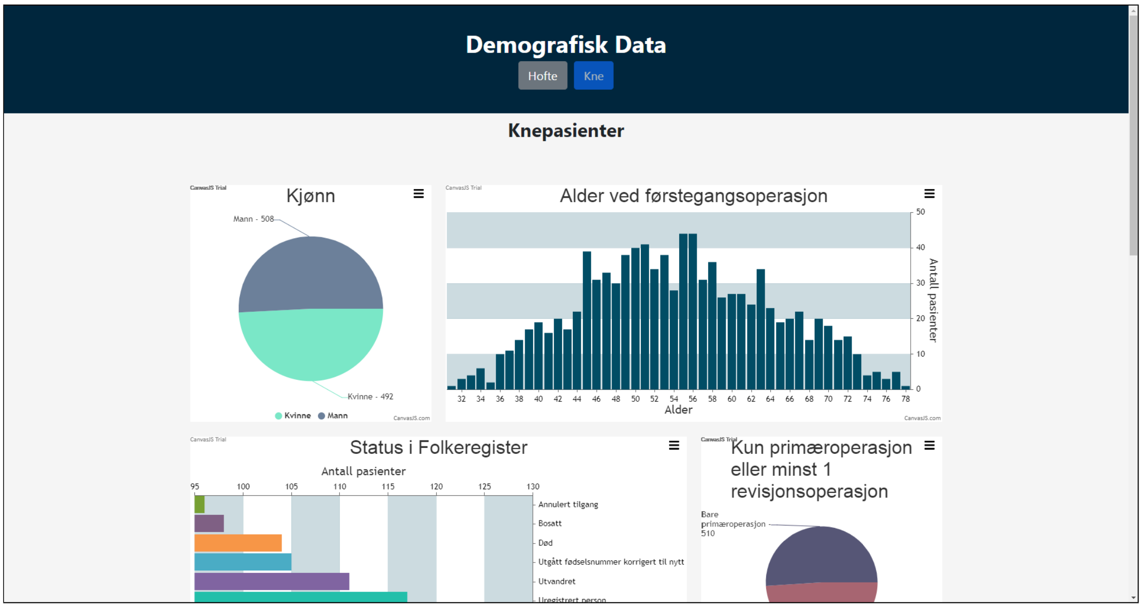


Figure 6.8: Demographic data page for knee arthroplasty patients

The entire page (edited together) of, in this case, knee patients' visualizations can be seen in Figure 6.9. The top-left graph shows the gender of the patients, the top right shows the age. The second row on the left shows the patient status in the national Resident Register, to the right of that shows whether patients have been operated on only once or several times. Row three shows a line graph of the number of primary operations over time separated by gender. The fourth row shows how long patients went between operations. The bottom graph displays five causes for revision surgery and the amount performed of each.

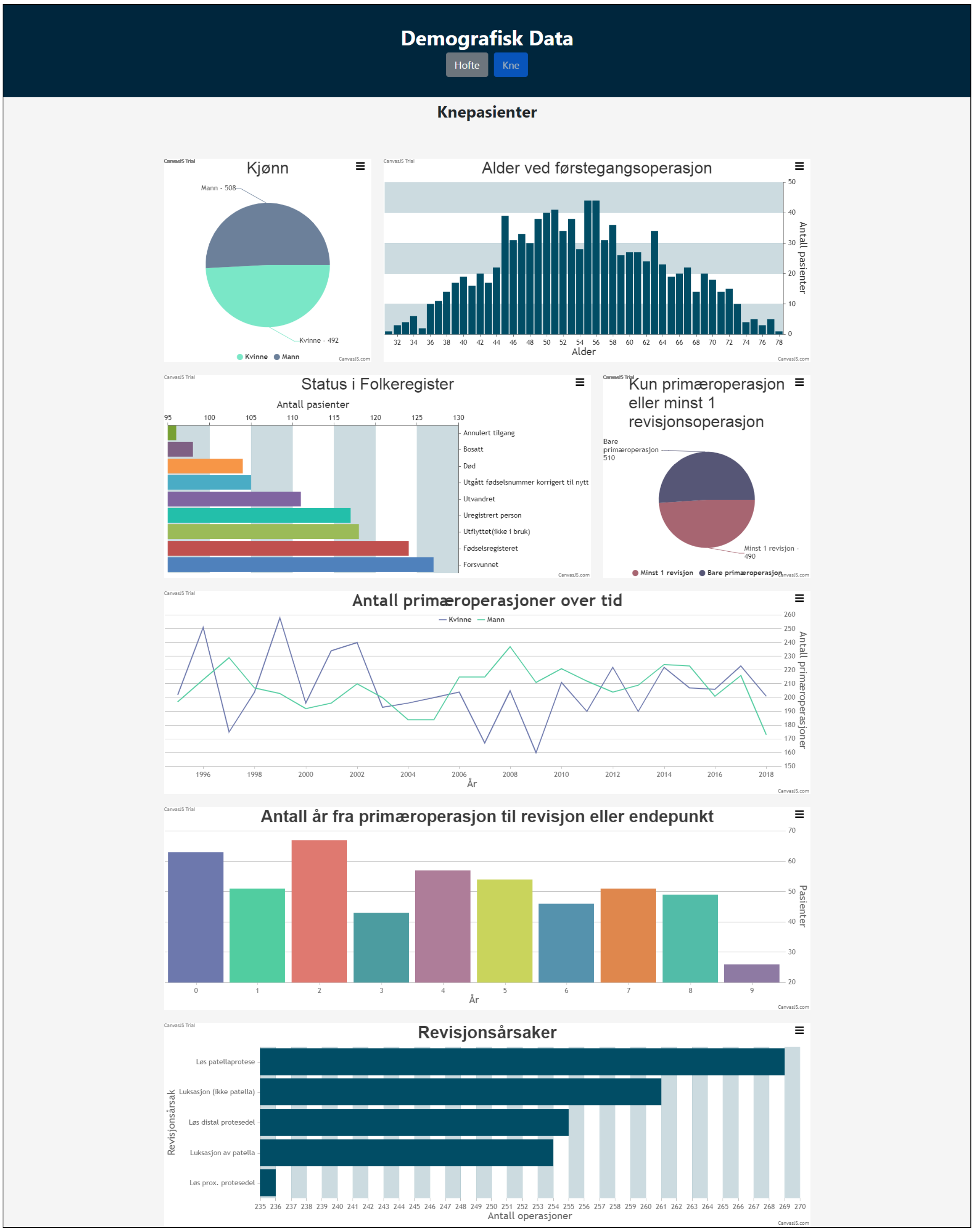


Figure 6.9: Full page for demographic data

Some of the graphs have the Y-axis labels showing on the right side. This was only due to a bug where the axis name did not show up if placed on the left side. As was eventually noted in the evaluation (specific comments are presented in Chapter 7), this is a somewhat unusual practice, but it was done purely due to this bug.

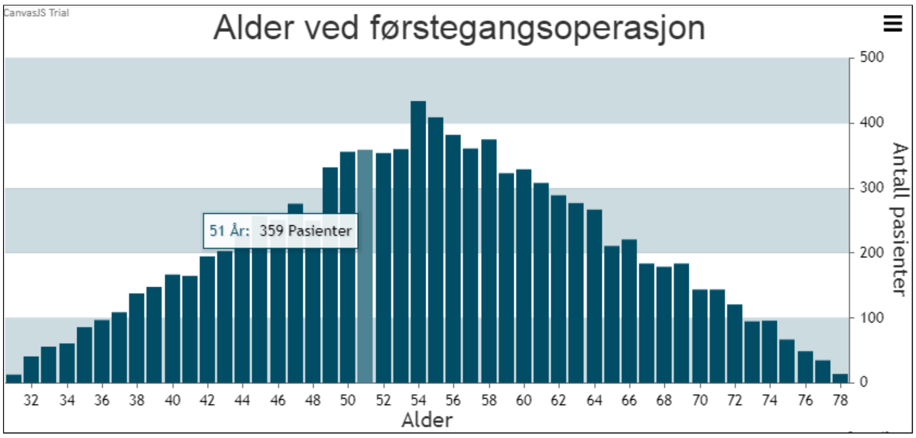


Figure 6.10: Demographic data graph interactivity

In Figure 6.10 one can see how data is presented when the user hovers the cursor over data on the graphs. This figure also shows the button that is present in the top right of each graph. This button gives the user different options for exporting image files of the chosen graph.

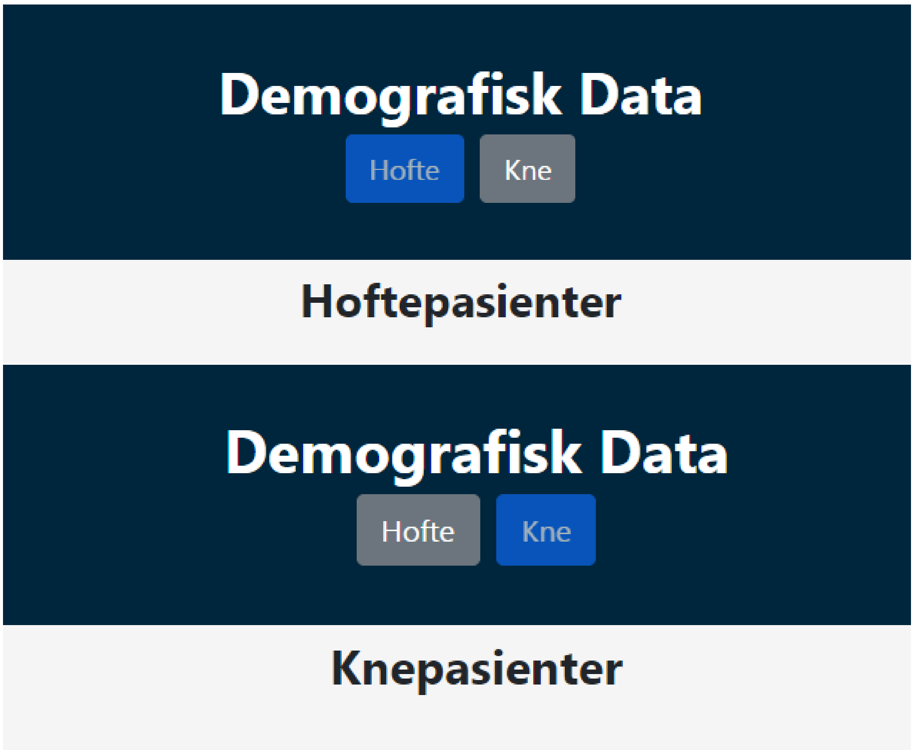


Figure 6.11: Different states of the demographic data buttons

Figure 6.11 shows the buttons that switch the data between hip and knee patients. When a button is pressed it changes to a blue color, the other button changes to gray, and the text underneath changes to show what data is active ('Hip patients' and 'Knee patients'). The activated button also becomes slightly dimmed and is not clickable while in this mode.

An additional graph was made that was intended to be implemented on the demographics page but was not due to time constraints. This graph can be seen in 6.12, 6.13, and 6.14. This graph showcases a functionality where the user can sort data by different age groups using the buttons underneath the graph. It also has the same interactive functionality as the other bar graphs on the demographic page.

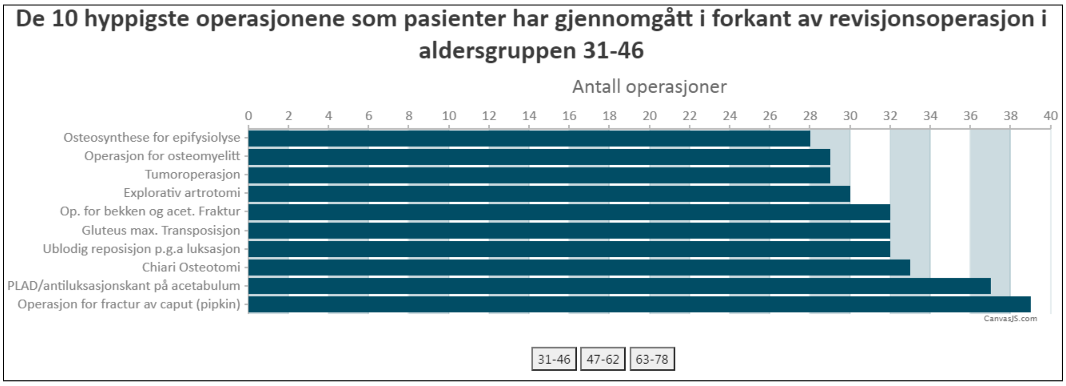


Figure 6.12: Demographic age group sorting 1

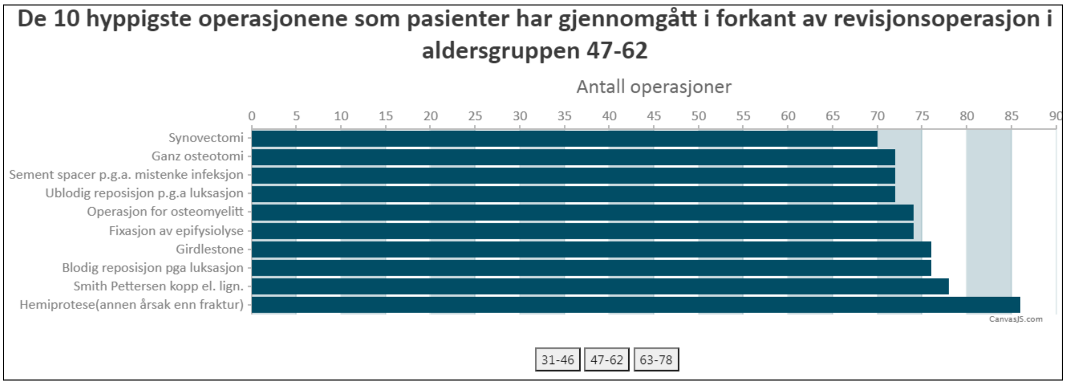


Figure 6.13: Demographic age group sorting 2

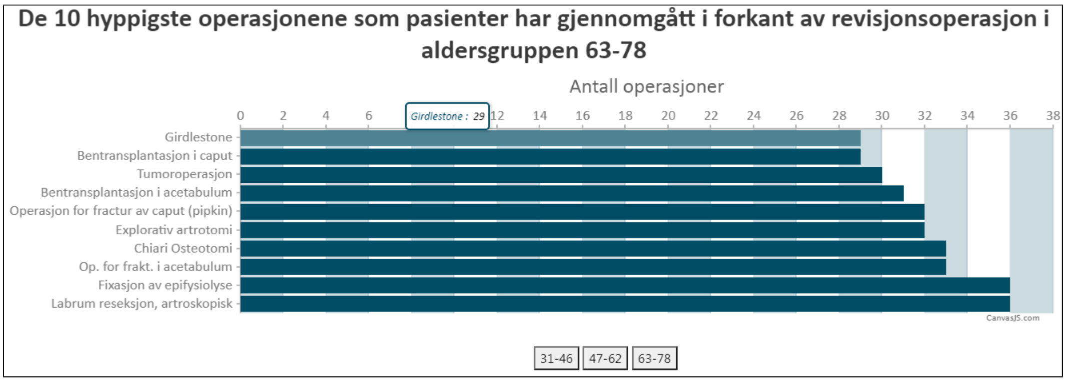


Figure 6.14: Demographic age group sorting 3

6.1.3 Kaplan-Meier

As the Kaplan-Meier graph is already being used in the register reports and by researchers in the field, it was a natural choice for visualizing for this project. In interviews with the register researchers, they indicated an interest in having the ability to go further in-depth in the data than just presenting one graph and table at a time. An interactive version of a Kaplan-Meier curve would allow a researcher to do exactly this. With functions like showing data when hovering over points on the graph, as well as zooming and panning in a graph, it would make it unnecessary to inspect tables of data to gain insight into the graph.

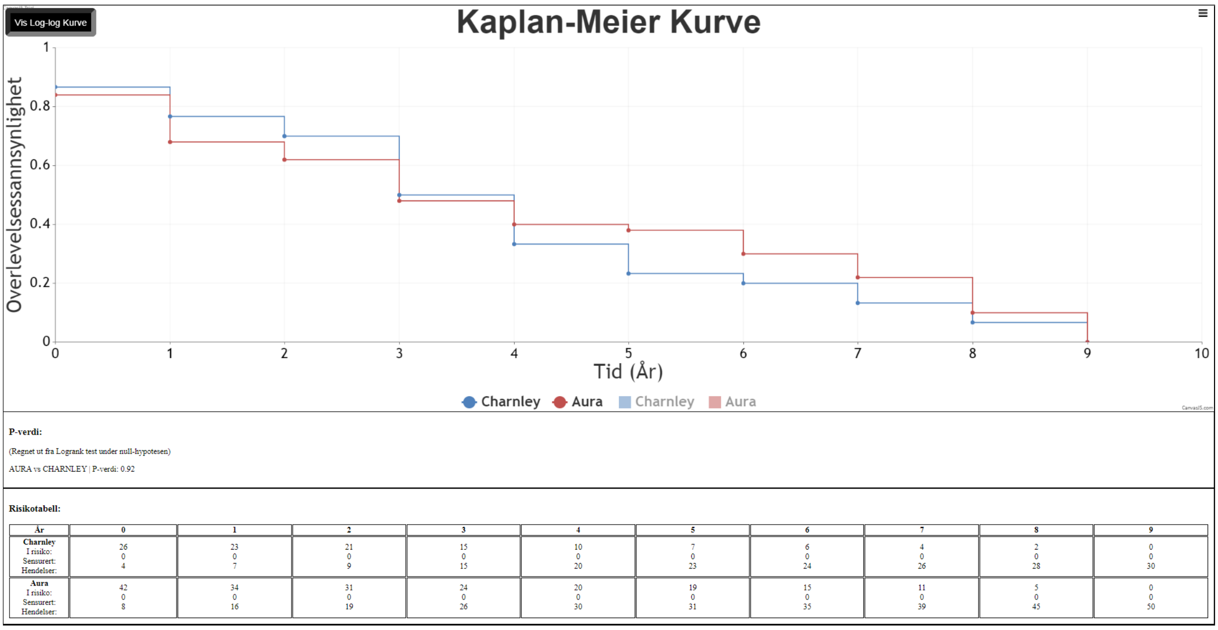


Figure 6.15: Kaplan-Meier graph

Figure 6.15 shows the graph with data from the back-end developers [24, 52] using CanvasJS. The graph shows the survival probability of different prostheses over a period of time. At the bottom of the page, a box shows the 'P values' of the different materials, an important detail for researchers looking at this type of data. The elements on the page have been slightly scaled down in order to show the entire page in the same image, in the actual version the user would have to scroll down in order to look at the table.

The colored circles and squares with names underneath the graph are clickable and change data currently being shown in the graph, allowing users to either compare different prostheses or focus on just one. Figure 6.16 shows what it looks like when only one prosthesis, in this case, 'Charnley', has been selected with these buttons. It also shows how the data is presented in the graph when hovering over one of the data points.

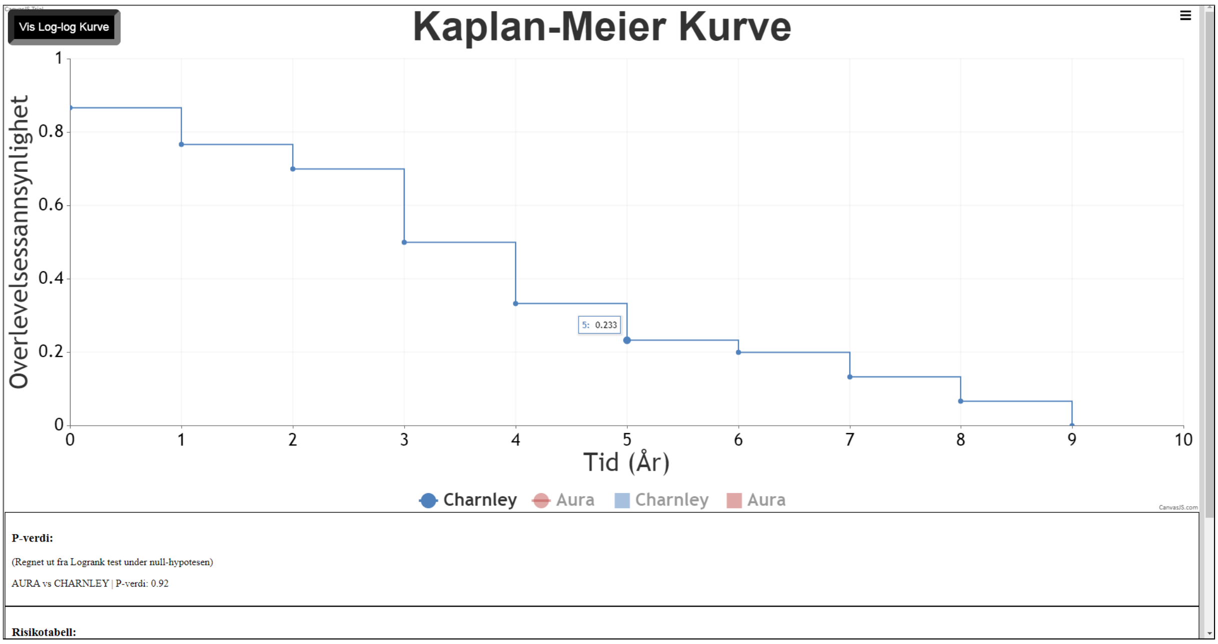


Figure 6.16: Kaplan-Meier button functionality

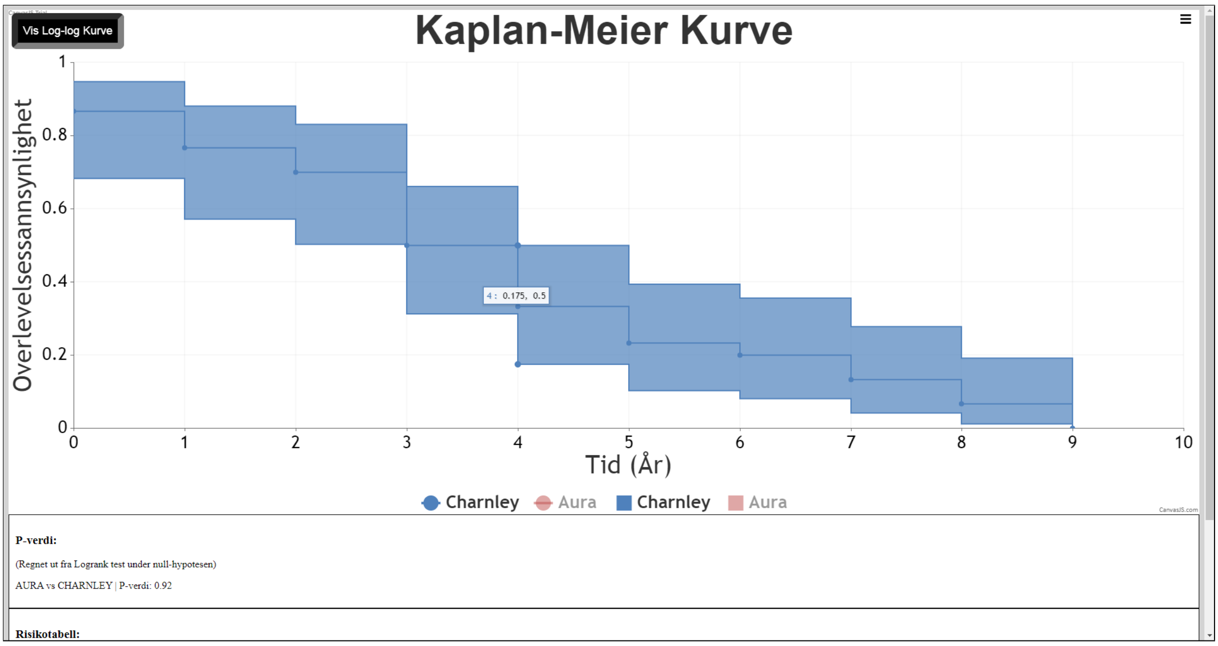


Figure 6.17: Kaplan-Meier survival estimate

In this case, the colored circles represent the survival probability for each material, while the colored squares represent the survival estimate (as discussed in Section 2.4.6). In Figure 6.17 the survival estimate for the selected material has been enabled on the graph. One can also see how data is presented when hovering.

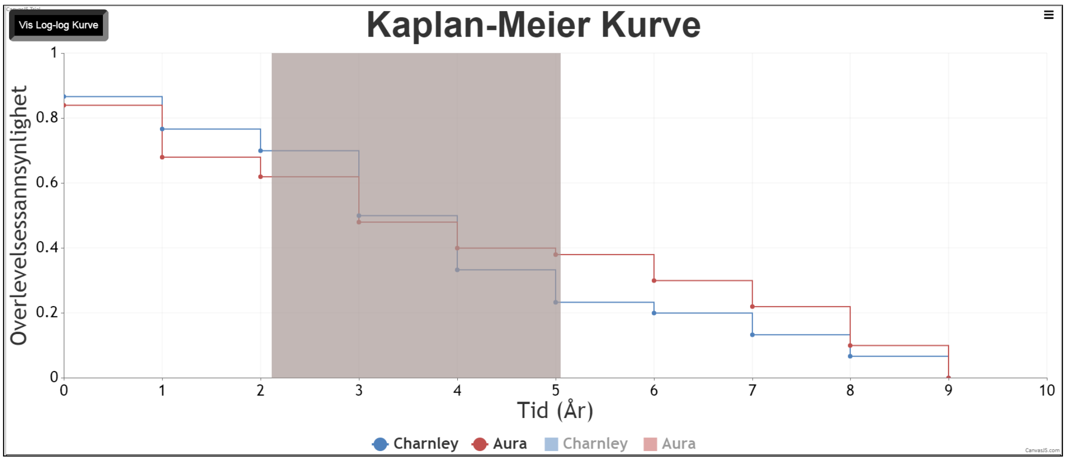


Figure 6.18: Kaplan-Meier Zooming selection

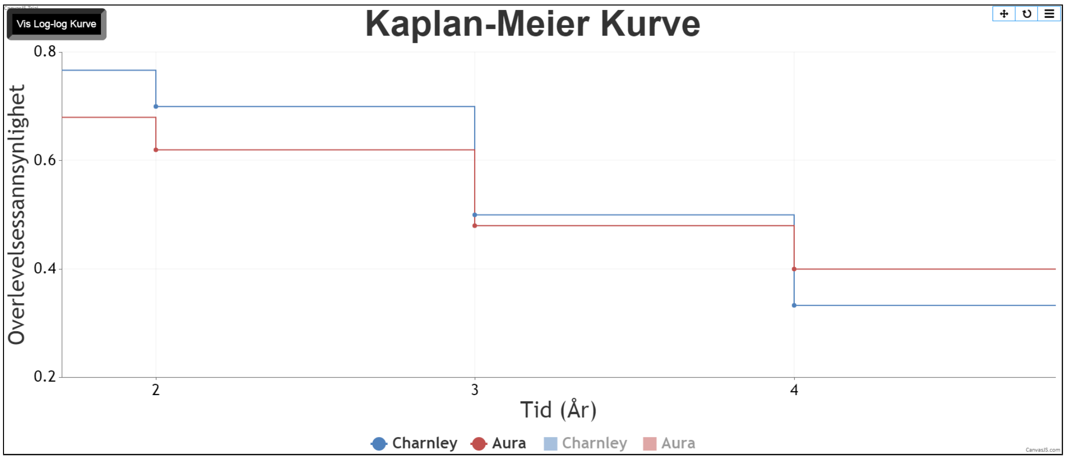


Figure 6.19: Kaplan-Meier Zoomed section

Figures 6.18 and 6.19 show the zooming functionality of the graph. The marked area in 6.18 shows what it looks like when the user clicks and drags over a selection of the graph which, once released, changes the graph to show only this section. When in this zoomed mode the user can click and drag to scan across different parts of the graph. Buttons for resetting the graph, or changing between zooming and panning, appear in the top right of the graph in this mode, as it can be seen in Figure 6.19.

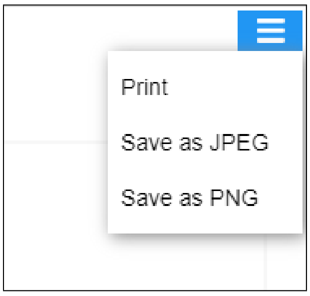


Figure 6.20: Kaplan-Meier Export function

In the top right of the graph is a button that allows the user to export the graph in different ways, shown in Figure 6.20.

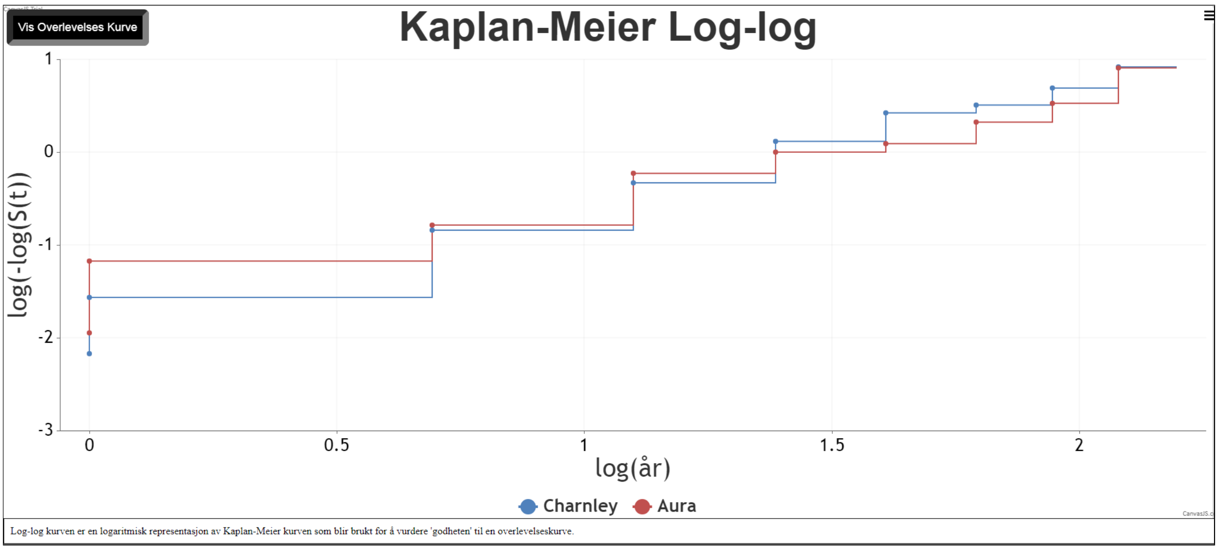


Figure 6.21: Kaplan-Meier Log-log plot

By pressing the button labeled 'Vis Log-log Kurve' (Show Log-log curve) in the top left of the graph it changes to a Log-log representation of the Kaplan-Meier graph. This can be seen in Figure 6.21. When viewing this graph the same button changes to 'Vis Overlevelse Kurve' (Show Survival Curve), which now brings the user back to the regular Kaplan-Meier curve when clicked. This graph has the same interactive functionality as shown on the regular graph, except the ability to show the confidence interval, as it is not used in this type of graph.

7 Evaluation

In several semi-structured interviews, we met with experts ranging from master graduates to biomedical researchers. This process was done in three parts; (i) a demonstration of the system open to questions or comments, (ii) a more in-depth discussion and question session after the demonstration, and finally (iii) a more detailed written feedback in three different forms that were to be filled out and returned after the interview. Some of the questions were specially created to address content in this project and were given together with a SUS form and Nielsen's hierarchy evaluation.

As each interview was providing feedback to all four students working on the register data, only the parts relevant for the visualizations are presented here. Three of these interviews were performed, with two to four experts participating at a time. The following sections will introduce the participants of the evaluation and the feedback they provided.

7.1 Participants

The following tables (7.1 and 7.2) present the participating experts and details on their expertise and experience. They are divided into domain experts and usability experts. Experts and their scores are presented separately to better understand their feedback.

Participants	Age	Gender	Education	Experience (years)
P1	41	Male	Ph.D.	10+
P2	45	Male	Ph.D.	20
P3	27	Male	M.Sc.	6
P4	29	Male	M.Sc.	5-10
P5	26	Male	M.Sc.	4

Table 7.1: Group 1: Domain experts

Participants	Age	Gender	Education	Experience (years)
P6	30	Female	M.Sc.	4
P7	32	Male	M.Sc.	4
P8	28	Male	M.Sc.	5

Table 7.2: Group 2: Usability experts

7.2 Feedback

The following are comments made during the interviews with the participants. Most comments are translated from Norwegian, and some comments are paraphrased for clarity.

7.2.1 First interview

County map

”That’s very beautiful, informative, but maybe it is good to have a short text on one side that says how it works [the visualization]. Does it have a tabular representation of the data? ” - P2

Demographic data

”Who is the target group, who is going to use this? Not all users would like to know if they’re ahead in age for operations, you need to consider their needs. But for researchers, this could be helpful because you have a lot of data where you can go into the details of it. Maybe some explanation and text could help to explain the data, for instance, a table alongside the data so you can use both information sources to highlight important information and visualize them as you hover over the table, for instance. So it’s important to know who you are designing for, who is the target group, how are you reaching that group, and why. Answers will come with time and user testing and so on.” - P8

Kaplan-Meier graph

On the log-log plot: ”It is good to represent it.” - P2

7.2.2 Second interview

County map

”I didn’t quite understand the year slider, I thought there were numbers missing. Maybe somehow show the range and make it clear it is a selection of years. Besides that, it is very nice and simple. It’s a very neat visualization, cool that you were able to make it. If the goal was to show the comparison function, it’s very clear. Maybe make the buttons bigger and more visible.” - P6

Demographic data

”Is there a reason why the y-axis labels are on the right side? Users usually read from left to right so it is more natural to have it on the left.” - P6 ”If a researcher wants to use the data it would be nice to be able to download the data in a format where they could be able to manipulate it further.” - P6

Kaplan-Meier graph

”When the confidence intervals are toggled it is a little hard to see the lines underneath them. Maybe increase the contrast in order to make it more readable. With all options selected it becomes hard to distinguish them.” - P6

7.2.3 Third interview

County map

”On the slider, you should show the beginning and end of the range to show you can go down to 1995 and up to 2018. There could also be a colormap displayed to show what the difference in color means on the map. It is not necessarily intuitively understood whether a bright or dark color is good or not. ” - P4

Demographic data

No comments were made.

Kaplan-Meier graph

No comments were made.

General Feedback

”The data visualizations were nice and had a good utilization of the data you had available.” - P7

7.3 Forms

The participants were given three forms to fill out in their own time after the interviews. Participant P2 was unable to submit the forms and has thus been excluded from the results.

7.3.1 System Usability Scale

Tables 7.3 and 7.4 show the calculated SUS score from both groups of evaluators.

Participant	Score
P1	90
P3	90
P4	82.5
P5	87.5

Table 7.3: Domain experts

Participant	Score
P6	60
P7	90
P8	85

Table 7.4: Usability experts

The average SUS score for the domain experts was 87.5 and for usability experts, it was 78.33. Figure 7.1 shows a column chart of the scores, displaying the different groups by color. Figure 7.2 shows the average of the scores divided by group.

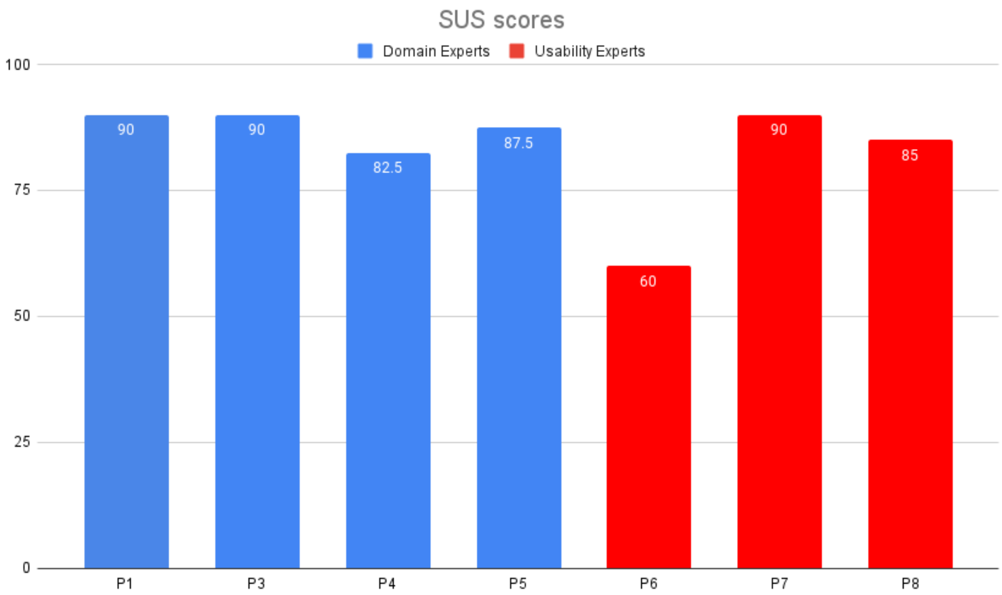


Figure 7.1: SUS scores by group

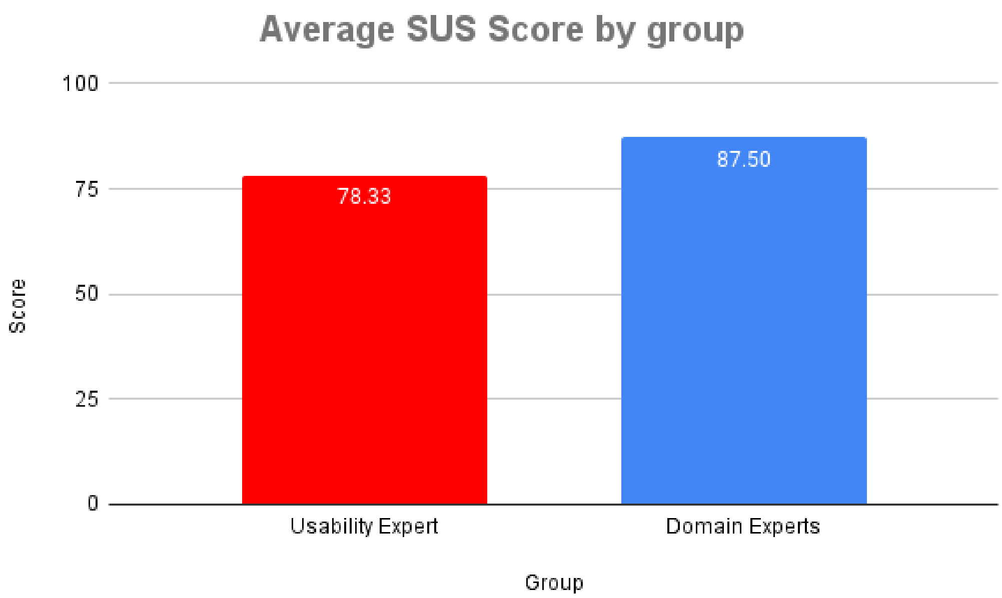


Figure 7.2: Average SUS scores by group

7.3.2 Nielsen Heuristics

Table 7.5 shows the average score for the heuristic evaluation for all participants.

Heuristic	**Average Score**
1. Visibility of system status	1
2. Match between system and the real world	0.833
3. User control and freedom	0.667
4. Consistency and standards	0.5
5. Error Prevention	0.333
6. Recognition rather than recall	0.167
7. Flexibility and efficiency of use	0.333
8. Aesthetic and minimalist design	0.167
9. Help users recognize, diagnose, and recover from errors	0.167
10. Help and documentation	0.667

Table 7.5: Results from Nielsen's heuristics form for both participant groups

The evaluators were also able to leave comments on the form. As the forms were filled out for all four students, the following comments are limited to those relevant for the visualizations. It is important to note that a lack of comments in these heuristics means a positive assessment. All comments that are given are intended to improve the current features.

Visibility of system status

One comment suggested improvement on the bar graphs by having numbers inside the bars in order to increase readability. Other comments recommended making a clear divide for the different user groups to avoid patient users 'stumbling upon' data meant for medical experts.

Match between system and the real world

Several comments once again emphasized the importance of separating the user groups. Language and visualizations should be adapted to fit different user groups.

User control and freedom

No relevant comments.

Consistency and standards

No relevant comments.

Error Prevention

The visualizations could benefit from some explanation, provide the user with instructions on how to operate them.

Recognition rather than recall

No comments.

Flexibility and efficiency of use

No comments.

Aesthetic and minimalist design

No relevant comments.

Help users recognize, diagnose, and recover from errors

No comments.

Help and documentation

A few comments suggested differentiating user groups and providing more instructions and explanations.

7.3.3 Additional visualization questions

Two additional questions were given to ten evaluators regarding the choice of visualization and level of interactivity using a Likert scale. One meant strong disagreement and five meant strong agreement. All feedback was positive; seven out of ten evaluators strongly agreed with the visualization (shown in Figure 7.3), five out of ten strongly agreed with the level of interactivity (shown in Figure 7.4).

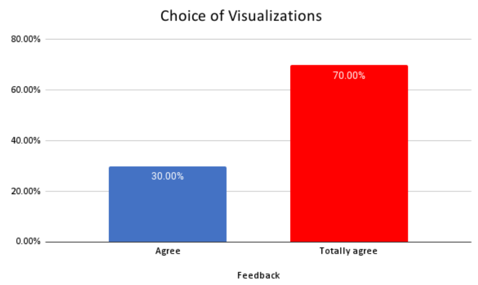


Figure 7.3: Choice of Visualizations feedback

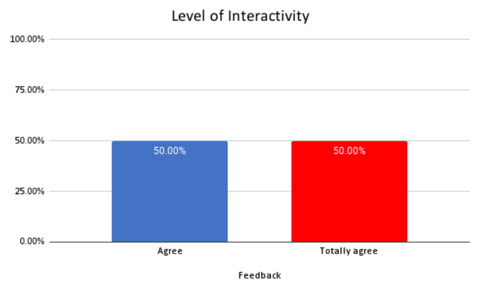


Figure 7.4: Level of Interactivity feedback

8 Discussion

This chapter discusses the methodologies and methods used in this research, the design, and development, as well as the main limitations. The three research questions are also answered here.

8.1 Design Science

The framework of Design Science was used throughout the project. Using the principles from this framework allowed for an efficient research process where the visualizations functioned as the artifacts. It was suitable for solving problems in the medical informatics domain as it has been demonstrated in several previous theses. Design Science is based on three pillars; Relevance, rigor, and design (Section 3.2.1).

Visualization and interactivity are relevant for the register and its users (Chapter 7). The design work has therefore been carried out, evaluated, and iterated through using methods that were most suitable for the problems at hand. The design has also happened as an interchange between back-end and front-end developers that worked on the project (Section 5.1), which meant lots of synergy in the team and discussion of various data mining methods, visualization, and HCI aspects.

8.2 Research Through Design (RTD)

Section 3.2.2 presented the concepts in RTD. One of these concepts speaks about contributions from research, where a design artifact can serve as one such contribution:

- If it solves a previously unsolved problem expands the knowledge base applies existing knowledge in innovative ways or produces a significant value the artifact itself can be seen as a contribution [51].

Results from the evaluation show that the visualizations have significant value and innovate on existing knowledge (Chapter 7). The visualizations also exemplify potentials in several of the other contributions in the RTD model: identifying new opportunities for technology advancement, providing researchers with inspiration and motivation for what they might build, identifying gaps in behavioral theory and models, and artifacts that provide a concrete embodiment of theory and technical opportunities which lead to practical applications of HCI research [51]. This last contribution is especially illustrated in the collaboration with the other students; using data and methods from the back-end data mining part of the project [24, 52] and the implementation of the visualizations in the HCI interface [49]. The collaboration also allowed us to focus research by not getting impeded by too many HCI considerations, which is another principle in the RTD model.

8.3 Visualization Development

The visualizations were developed through four iterations (Chapter 5). The reason was that each visualization had followed the development of the back-end part. This also influenced the different types of visualizations produced, such as the county map, demographic data, and Kaplan-Meier (Chapter 6). The team discussions were used to iterate through different graphical details, choice of colors in graphs, amount of content in graphs, and helped capture different opinions that could be expected in user groups. The solutions appearing in the final visualizations are representative of both the development team and the evaluators but could be further fine-tuned and modified in future iterations.

The demonstration and interaction with visualizations have been instrumental during both development and testing. Some visualizations were close to the methods and therefore will still be very meaningful to the experts only, while the county map seemed to be popular and easy to explore, as seen during the evaluation (Chapter 7).

8.4 Evaluation of Visualizations

Visualizations were tested as a part of the front-end system, which was demonstrated as a whole. Extra time was taken to go into some visualization details. This way the evaluators could receive a feeling of the purpose of the visualization in a wider context. The methods are not standing only by themselves in reality, even though they were developed independently of the user interface, which was the focus of another master's project [49]. It has to be said that the users have experienced the whole front-end system as one, and they gave scores and comments for the entire front-end solution. The choice of visualization, however, was addressed separately and found to be highly acceptable as well as the interactivity aspect of it. Some evaluators found that additional visualizations could be added (Section 7.3.3).

8.4.1 SUS

Positive SUS results of the front-end suggest an agreement of the evaluators with visualization and suitability of the artifact within the arthroplasty domain. The domain experts gave the system an average score of 87.5. Some issues were highlighted by the usability experts who gave the system a somewhat lower score of 78.33 (Section 7.3.1). Since it is established in the SUS literature that a score of more than 68 is considered a good score [34], these scores are above the required threshold.

8.4.2 Nielsen's Heuristics

Experts have pointed out several aspects of the system that could be made better and even suggested some practical improvements. For example, some comments regarding explanatory text in graphs should be differently positioned to help users or making certain color contrasts clearer. Another important suggestion was to separate user groups and provide additional information in form of help and instructions (Section 7.3.2).

A total of ten evaluators provided comments that the development team could not foresee, so it has proven to be useful to perform this kind of intuitive and simple expert evaluation.

8.5 Limitations

There were several limitations encountered during the project. Many of these were brought on by the lock-downs due to the COVID-19 pandemic. There was a delay in gaining access to the data we were to use in the project because of complications in communications with the register, so finally, we settled on generated sample data instead of the actual data. Due to this delay and the lack of real data, the number of visualizations produced was somewhat limited. Meetings, interviews, and evaluations had to be done remotely, and access to different user groups such as patients and physicians became severely restricted. Lack of access to the University campus created additional limitations on resources provided by the University.

8.6 Answering Research Question

8.6.1 RQ1: How can data visualization help doctors and patients get a better understanding of medical data?

The research in this thesis has shown a great acceptance of the visualizations proposed and demonstrated during the evaluation process. All evaluators were agreeable that they liked the choice of methods and level of interactivity (Chapter 7). The majority of the comments were very positive, encouraging, and most of the suggestions were only for minor improvements.

The long-term effects of the visualizations and interactivity on the register data are yet to be explored. The literature illustrates good results that show improvement in relationships between users, patients, and solution developers. Section 2.4.5 is suggesting improvement in reporting of interventions and working with patient registers. These publications are very encouraging and indicate that efficient visualization could indeed impact the clinical practice positively in many ways.

8.6.2 RQ2: Can arthroplasty data from the register be visualized in a more efficient and informative way to meet different user groups' needs?

Some of the earlier master's theses (presented in Section 2.6) have given a number of excellent examples of visualizations, from basic demographic information to the presentation of results from principal component analysis and cluster analysis carried out on the register data. There was also one user interface developed on the related set of data to enable easier work and presentation of the results. Evaluation in all these theses was very positive and has inspired further work presented in this thesis (Chapter 7). The visualization methods and interaction have been produced independently of the framework or system in which they could be implemented in order to focus on these two aspects, using data for both knee and hip prostheses [24, 52]. The purpose was also to develop and suggest methods that could function with various types of data with only minor adjustments. Since the register is currently using only statistical methods and presenting data in basic tables and charts [20], there is potential for more visual and interactive ways to enable users to perform their own data exploration using various data mining methods.

The current set of visualization methods was met with positive responses from the majority of evaluators (Chapter 7). It seems to provide lots of functionality but could be extended to include some more methods. These could be done in cooperation with different user groups, but in the first place with physicians and biomedical researchers interested in working with the data.

Patients are a group to consider in further analysis of requirements and evaluation (Section 7.3.2). It could be worthwhile to make efforts towards other non-expert groups since the register is already offering some information to the public through the annual reports [20]. By looking at the reports in their current form, it is clear that they are suited more for experts than the general public. There is also potential that goes beyond evaluating and looking at just the visualization methods. The literature suggests improving healthcare by introducing more contemporary visualization methods (Section 2.4.5).

8.6.3 RQ3: Should sensitive medical data be presented to patients and be open to the public?

There are many definitions of ’sensitive’ medical data. The most intuitive understanding of this term has to do with sharing very personal data such as patient name, age, location, gender, etc. However, there are many other understandings of ’sensitive’ that have to do with the interpretation of the data and even methods used to analyze the data. During the evaluation, some participants had concerns about the meaning of ’survival analysis’ since it could be misinterpreted by non-expert users as regarding the survival of patients, not the prostheses (Chapter 7). This suggested one should be cautious with what to present to patients and the public. Similarly, presenting comparisons of different kinds of prostheses might be misinterpreted if it looks from a graph that one kind of implant is objectively ’better’ than another. The success of surgery depends on many additional factors. Thus results have to be interpreted correctly, as they are often complex, and should perhaps be left only to medical experts and researchers.

Seeing demographic data could be a good type of information for non-experts. A patient might see that there are others who got implants at a similar age as them and that there are common and acceptable solutions for securing a better quality of life, even through surgery. Graphical presentations of implants could be appreciated by patients interested in learning more about the procedure and implants that they will receive. This is another field of research where it will be important to include a number of patients of different ages, gender, background, etc. in order to understand particular preferences and needs. Visualization is there to help deliver the information in a more appealing, informative, and interactive way, however, it has to be the appropriate information.

9 Conclusions and Future Work

By following the Design Science methodology we have ensured relevance, rigor, and quality that resulted in an artifact that is comprised of several types of visualizations with interactive functionality. All work was done using both knee and hip prosthesis data and has shown through the evaluation that it had appeal and met certain user expectations. Since the arthroplasty register is employing traditional ways of working with data using multivariate statistics and published reports and scientific papers, the contribution of this thesis could be deemed as novel as it brought a whole new set of possibilities to visualize and interact with data.

Visualization and modeling of the data could be seen as an independent cycle of development, but it also has been a part of the bigger project in which the selection of data analytical methods was suggested by the back-end team development. Moreover, the work of this thesis was part of a front-end development in which a web-based high-fidelity prototype was produced. Through this prototype, it became clear that the visualization was made easy for the final users as they could choose data analysis of the implants they were most interested in. By doing so they were guided to the visualization of the data in which they could select what period of data to look at, demographics of the population, what are the differences in longevity between two prostheses, etc. Compared to the work in previous master's theses, this was a significant step towards the users. However, the greater part of the system is still directed towards the expert users. Parts of the system that could accommodate patients and other users from the public are underrepresented, but could surely be expanded once we know what data analysis and results can be shared with a broader audience. Restrictions for sharing are not limited just to secure and less secure data, but also to the current lack of proper functionality that would help to interpret their results.

Regardless of the limitations of the current situation, a solid evaluation was conducted using domain and IT experts. Obtained usability scores are encouraging and many comments were made regarding the usefulness of visualization and interactivity. The utility of the functionality that was presented was only enhanced by the choice of data mining methods and the web-based system which can be only attributed to the way that the project has been carried out. A close connection of the back-end and front-end development has enabled a realistic experience and helped deliver results to the extent that has not been seen with the register data by now. For example, a set of 'dry' tables and calculations was also equipped with visualization that could be further explored by clicking and selecting a preferred time period or part of the county in which the user was interested. Using a map of Norway was a novel feature, and was much appreciated and accepted. One functionality enabled comparing the top 5 prostheses in different counties. Such a comparison usually takes more time, since the data has to be searched for in different tables. Having interactive solutions could help the performance of tasks like this in smoother ways and help physicians and administration save time and effort.

9.1 Future work

The development team ought to include in the future more physicians, researchers, and a diverse patient group. It is especially for patients and the public that the current system is not fully utilized, so the needs of these groups remain to be studied in depth. Interpretation of the results seems not to be critical at this stage, but as the user groups will expand this is one of the functionalities that will need more attention.

Current forms of visualizations are appreciated by those who evaluated them, but there are many more possibilities to use different graphs, colors, and create output according to information needs and user preferences that will be identified in the future. The amount of information and visualizations are rather high for expert users, but there are other forms that patients would appreciate. Using the map of Norway could be done in several stereotypical ways, where the data could be selected to suit different user groups. For example, management might be interested in all the details of the prostheses for billing and similar administrative purposes while it could be expected that surgeons would be more interested in outcomes and the main prosthesis features they are working with.

Maintaining the visualization functionalities will depend on the data mining methods on the back-end as well the interface on the front-end side. Both these sides are expected to evolve and include new methods and new procedures as well as user groups which will also demand that visualization is maintained and upgraded accordingly.

For this project, it was necessary to work on parts of the system separately, but once the prototype is in use and further developed, the visualization aspect will have to deal with safety, user-friendliness, and include far more features for patients and the public. There is a place to develop a more user-centered application throughout the iterations that will also be based on new user feedback. One could start even by implementing a standard disease-related questionnaire like SF-36 and invite patients to visualize their results using values of their own scores entered pre- and post-operatively, and even during a longer period of rehabilitation. Using a web application and visualizing results would be of appeal as compared to filling in papers that have to be entered by somebody else and presented with a time delay. Implementing standard tools such as SF-36 could also help the register understand data coming from patients based on their self-monitoring.

It is expected, due to the literature review, that visualization could have a high clinical utility contributing to decision-support, calculation of outcomes, identifying critical features, and helping build partnerships between healthcare providers and patients over efficient visualizations. This is one future direction to study in a real clinical environment.

References

[1] Ali, S. M., N. Gupta, G. K. Nayak, and R. K. Lenka (2016). Big data visualization: Tools and challenges. In *Proceedings of the 2016 2nd International Conference on Contemporary Computing and Informatics, IC3I 2016*, pp. 656–660. Institute of Electrical and Electronics Engineers Inc. 1, 12

[2] Atlassian (2021). About | what is trello? https://trello.com/about. [Online; accessed 11-June-2021]. 32

[3] Bellazzi, R. and B. Zupan (2008, February). Predictive data mining in clinical medicine: Current issues and guidelines. 6, 7

[4] Bostock, M. (2020). D3.js - data driven documents. https://d3js.org/. [Online; accessed 09-May-2020]. 38

[5] Chou, L. (2019). 9 data visualization tools that you cannot miss in 2019. https://towardsdatascience.com/9-data-visualization-tools-that-you-cannot-miss-in-2019-3ff23222a927. [Online; accessed 12-June-2021]. 7

[6] Clifton, C. (2017). Data mining | computer science | Britannica.com. https://www.britannica.com/technology/data-mining. [Online; accessed 11-June-2021]. 1

[7] CodePen (2021). Codepen. https://codepen.io/. [Online; accessed 09-June-2021]. 38

[8] Corporation, M. (2021). Visual studio code - code editing. redefined. https://code.visualstudio.com/. [Online; accessed 11-June-2021]. 33

[9] Crampton, J. W. (2002). Interactivity types in geographic visualization. *Cartography and Geographic Information Science 29*(2), 85–98. 27

[10] Davis Giardina, T., S. Menon, D. E. Parrish, D. F. Sittig, and H. Singh (2014, jul). Patient access to medical records and healthcare outcomes: a systematic review. *Journal of the American Medical Informatics Association 21*(4), 737–741. 16, 17

[11] Dimara, E. and C. Perin (2020, jan). What is Interaction for Data Visualization? *IEEE Transactions on Visualization and Computer Graphics 26*(1), 119–129. 8, 9, 10

[12] Dresch, A., D. P. Lacerda, and J. A. V. Antunes (2015). *Design Science Research: A Method for Science and Technology Advancement*. Springer. 22

[13] Drolet, B. C. and K. B. Johnson (2008, dec). Categorizing the world of registries. *Journal of Biomedical Informatics 41*(6), 1009–1020. 4

[14] Elliott, D., R. Lazarus, and S. R. Leeder (2006, jul). Health outcomes of patients undergoing cardiac surgery: Repeated measures using Short Form-36 and 15 Dimensions of Quality of Life questionnaire. *Heart and Lung: Journal of Acute and Critical Care 35*(4), 245–251. 15, 21

[15] Faiola, A., P. Srinivas, and J. Duke (2015). Supporting Clinical Cognition: A Human-Centered Approach to a Novel ICU Information Visualization Dashboard. *AMIA ... Annual Symposium proceedings. AMIA Symposium 2015*, 560–569. 13

[16] Fenopix (2021). About us | canvasjs. `https://canvasjs.com/about/`. [Online; accessed 11-June-2021]. 32

[17] Foundation, P. S. (2021). About python™ | python.org. `https://www.python.org/`. [Online; accessed 11-June-2021]. 32

[18] Franklin, A., S. Gantela, S. Shifarraw, T. R. Johnson, D. J. Robinson, B. R. King, A. M. Mehta, C. L. Maddow, N. R. Hoot, V. Nguyen, A. Rubio, J. Zhang, and N. G. Okafor (2017, july). Dashboard visualizations: Supporting real-time throughput decision-making. *Journal of Biomedical Informatics 71*, 211–221. 13

[19] Hand, D. J. (2007, nov). Principles of data mining. In *Drug Safety*, Volume 30, pp. 621–622. Springer. 4

[20] Haukeland, U. H. (2020). Norwegian national advisory unit on arthroplasty and hip fractures. `http://nrlweb.ihelse.net/eng/`. [Online; accessed 06-May-2020]. 4, 21, 66

[21] Havelin, L. I., L. B. Engesaeter, B. Espehaug, O. Furnes, S. A. Lie, and S. E. Vollset (2000, jan). The Norwegian Arthroplasty Register: 11 years and 73,000 arthroplasties. *Acta Orthopaedica Scandinavica 71*(4), 337–353. 1

[22] Hearst, M. (2011). User interfaces for search. *Modern Information Retrieval*, 21–55. 12, 13

[23] Hevner, A. R., S. T. March, J. Park, and S. Ram (2004, mar). Design science in information systems research. *MIS Quarterly: Management Information Systems 28*(1), 75–105. 22, 23, 24

[24] Hufthammer, K. T. (2021). Data mining for outcome analysis in hip arthroplasty. Master's thesis, University of Bergen. 32, 39, 40, 43, 53, 63, 65

[25] Iden, A. (2020). Data mining approach to modelling of outcomes in total knee arthroplasty. Master's thesis, University of Bergen. 19, 27

[26] Inc., G. (2020). Google forms: Free online surveys for personal use. `https://www.google.com/forms/about/`. [Online; accessed 08-May-2020]. 29

[27] Kanbanize (2021). Kanban explained for beginners | the complete guide. `https://kanbanize.com/kanban-resources/getting-started/what-is-kanban`. [Online; accessed 11-June-2021]. 26

[28] Kelleher, C. and T. Wagener (2011, jun). Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling and Software 26*(6), 822–827. 1, 10, 12

[29] Khairat, S. S., A. Dukkipati, H. A. Lauria, T. Bice, D. Travers, and S. S. Carson (2018, may). The Impact of Visualization Dashboards on Quality of Care and Clinician Satisfaction: Integrative Literature Review. *JMIR Human Factors 5*(2), e22. 13

[30] Kim, J. (2019, mar). Drawing guideline for JKMS manuscript (01) Kaplan-Meier curve and survival analysis. *Journal of Korean Medical Science 34*(8). 1

[31] Kishore, J., M. Goel, and P. Khanna (2010). Understanding survival analysis: Kaplan-Meier estimate. *International Journal of Ayurveda Research 1*(4), 274. 13, 14

[32] Kristoffersen, Y. (2019). Mining for individual patient outcome prediction in hip arthroplasty registry data. Master's thesis, University of Bergen. 20, 27

[33] Lazar, J., J. H. Feng, and H. Hochheiser (2017). *Research methods in human-computer interaction* (2 ed.). Morgan Kaufmann Publishers, an imprint of Elsevier. 28, 29

[34] Lewis, J. R. (2018, jul). The System Usability Scale: Past, Present, and Future. *International Journal of Human-Computer Interaction 34*(7), 577–590. 30, 64

[35] Longberg, P.-N. (2019). Hale, the hip arthroplasty longevity estimation system. Master's thesis, University of Bergen. 17, 19

[36] Mariconda, M., O. Galasso, G. G. Costa, P. Recano, and S. Cerbasi (2011, dec). Quality of life and functionality after total hip arthroplasty: A long-term follow-up study. *BMC Musculoskeletal Disorders 12*(1), 222. 16

[37] Medicine, J. H. (2021). Arthroplasty. `https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/arthroplasty`. [Online; accessed 13-April-2021]. 3

[38] Mozilla (2021). Web technology for developers. `https://developer.mozilla.org/en-US/docs/Web/JavaScript`. [Online; accessed 19-April-2021]. 32

[39] Mugisha, A., P. Wakholi, and A. Babic (2019). Design features for usable mobile electronic data capturing forms: The form developers' perspective. In L. Lhotska, L. Sukupova, I. Lacković, and G. S. Ibbott (Eds.), *World Congress on Medical Physics and Biomedical Engineering 2018*, Singapore, pp. 463–466. Springer Singapore. 27

[40] Nielsen, J. (1994a). 10 usability heuristics for user interface design. `https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/`. [Online; accessed 07-June-2021]. 30

[41] Nielsen, J. (1994b). How to conduct a heuristic evaluation. `https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/`. [Online; accessed 07-June-2021]. 30

[42] Osborne, J. (2014, jan). *Best Practices in Data Cleaning: A Complete Guide to Everything You Need to Do Before and After Collecting Your Data*. SAGE Publications, Inc. 29

[43] Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. (2011). Scikit-learn: Machine learning in python. *the Journal of machine Learning research 12*, 2825–2830. 26, 32

[44] Project, T. I. (2021). About | inkscape. `https://inkscape.org/about/`. [Online; accessed 11-June-2021]. 33, 40

[45] Rich, J. T., J. G. Neely, R. C. Paniello, C. C. Voelker, B. Nussenbaum, and E. W. Wang (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngology - Head and Neck Surgery 143*(3), 331–336. 1, 14

[46] Rivard, S., G. Poirier, L. Raymond, and F. Bergeron (1997, jun). Development of a Measure to Assess the Quality of User- Developed Applications. *Data Base for Advances in Information Systems 28*(3), 44–58. 27

[47] Shneiderman, B., C. Plaisant, and B. W. Hesse (2013). Improving healthcare with interactive visualization. *Computer 46*(5), 58–66. 1

[48] SimpleMaps (2021). Simplemaps - interactive maps & data. `https://simplemaps.com/`. [Online; accessed 09-June-2021]. 38

[49] Stolt-Nielsen, S. B. (2021). Design driven development of a web-enabled system for data mining in arthroplasty registry. Master's thesis, University of Bergen. 34, 36, 42, 63, 64

[50] Zhang, S., C. Zhang, and Q. Yang (2003, may). Data preparation for data mining. *Applied Artificial Intelligence 17*(5-6), 375–381. 5, 6, 26

[51] Zimmerman, J., J. Forlizzi, and S. Evenson (2007). Research through design as a method for interaction design research in hci. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 493–502. 24, 25, 28, 63

[52] Ånneland, S. (2021). Web-based data mining tool for total knee arthroplasty. Master's thesis, University of Bergen. 32, 39, 40, 43, 53, 63, 65

A Appendix - NSD Approval

NORSK SENTER FOR FORSKNINGSDATA

NSD sin vurdering

Prosjekttittel

Arthroplasty data mining and visualization

Referansenummer

219665

Registrert

20.10.2020 av Arle Farsund Solheim - Arle.Solheim@student.uib.no

Behandlingsansvarlig institusjon

Universitetet i Bergen / Det samfunnsvitenskapelige fakultet / Institutt for informasjons- og medievitenskap

Prosjektansvarlig (vitenskapelig ansatt/veileder eller stipendiat)

Ankica Babic, Ankica.Babic@uib.no, tlf: 4755589139

Type prosjekt

Studentprosjekt, masterstudium

Kontaktinformasjon, student

Arle Farsund Solheim, arle.farsund.solheim@gmail.com, tlf: xxxxxxxx

Prosjektperiode

15.11.2020 - 15.06.2021

Status

06.11.2020 - Vurdert

Vurdering (1)

06.11.2020 - Vurdert

Det er vår vurdering at behandlingen av personopplysninger i prosjektet vil være i samsvar med personvernlovgivningen så fremt den gjennomføres i tråd med det som er dokumentert i meldeskjemaet den 06.11.2020 med vedlegg, samt i meldingsdialogen mellom innmelder og NSD. Behandlingen kan starte.

MELD VESENTLIGE ENDRINGER
Dersom det skjer vesentlige endringer i behandlingen av personopplysninger, kan det være nødvendig å melde dette til NSD ved å oppdatere meldeskjemaet. Før du melder inn en endring, oppfordrer vi deg til å lese om hvilke type endringer det er nødvendig å

melde: https://nsd.no/personvernombud/meld_prosjekt/meld_endringer.html
Du må vente på svar fra NSD før endringen gjennomføres.

TYPE OPPLYSNINGER OG VARIGHET
Prosjektet behandler særlige kategorier av personopplysninger om helse og alminnelige kategorier av personopplysninger frem til 15.06.2021.

LOVLIG GRUNNLAG
Prosjektet vil innhente samtykke fra de registrerte til behandlingen av personopplysninger. Vår vurdering er at prosjektet legger opp til et samtykke i samsvar med kravene i art. 4 nr. 11 og art. 7, ved at det er en frivillig, spesifikk, informert og utvetydig bekreftelse, som kan dokumenteres, og som den registrerte kan trekke tilbake. Lovlig grunnlag for behandlingen er dermed den registrertes uttrykkelige samtykke, jf. personvernforordningen art. 6 nr. 1 bokstav a, jf. art. 9 nr. 2 bokstav a, jf. personopplysningsloven § 10, jf. § 9 (2).

PERSONVERNPRINSIPPER
NSD vurderer at den planlagte behandlingen av personopplysninger vil følge prinsippene i personvernforordningen om:

- lovlighet, rettferdighet og åpenhet (art. 5.1 a), ved at de registrerte får tilfredsstillende informasjon om og samtykker til behandlingen

- formålsbegrensning (art. 5.1 b), ved at personopplysninger samles inn for spesifikke, uttrykkelig angitte og berettigede formål, og ikke viderebehandles til nye uforenlige formål

- dataminimering (art. 5.1 c), ved at det kun behandles opplysninger som er adekvate, relevante og nødvendige for formålet med prosjektet

- lagringsbegrensning (art. 5.1 e), ved at personopplysningene ikke lagres lengre enn nødvendig for å oppfylle formålet

DE REGISTRERTES RETTIGHETER
Så lenge de registrerte kan identifiseres i datamaterialet vil de ha følgende rettigheter: åpenhet (art. 12), informasjon (art. 13), innsyn (art. 15), retting (art. 16), sletting (art. 17), begrensning (art. 18), underretning (art. 19), dataportabilitet (art. 20).

NSD vurderer at informasjonen som de registrerte vil motta oppfyller lovens krav til form og innhold, jf. art. 12.1 og art. 13.
Vi minner om at hvis en registrert tar kontakt om sine rettigheter, har behandlingsansvarlig institusjon plikt til å svare innen en måned.

FØLG DIN INSTITUSJONS RETNINGSLINJER
NSD legger til grunn at behandlingen oppfyller kravene i personvernforordningen om riktighet (art. 5.1 d), integritet og konfidensialitet (art. 5.1. f) og sikkerhet (art. 32).

For å forsikre dere om at kravene oppfylles, må dere følge interne retningslinjer og eventuelt rådføre dere med behandlingsansvarlig institusjon.

OPPFØLGING AV PROSJEKTET
NSD vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til med prosjektet!

Kontaktperson hos NSD: Henrik Netland Svensen
Tlf. Personverntjenester: 55 58 21 17 (tast 1)

B Appendix - Interview Guides and Consent form

Intervjuguide

Format: ansikt til ansikt eller digitalt intervju.

Svarregistrering: Lydopptaker, notater.

Hensikten med intervjuet er å først etablerte den faglige bakgrunn hos deltakerne, la deltakerne utforske og tolke resultatene og modellene jeg har utviklet, for deretter å få deltakernes oppfattelse av resultatene og hvorvidt de er aktuelle for fremtidige informasjonssystemer i ortopedi.

Intervju

Varighet rundt 30-60 minutter.

Deltakerne informeres om hva prosjektet går ut på og hva jeg ønsker med intervjuet. Deltakernes teknologiske kompetanse og deres kunnskap rundt artroplastikk og databaser kartlegges.

Spørsmål:

- Hvilken erfaring har du med teknologi?
    - Vil du si du er datakyndig (kompetent)?
- Hvilken teknologier bruker du på en daglig basis?
- Kan du fortelle om en vanlig dag for deg?
    - (hvor teknologi bli brukt)

- Hvordan vil du beskrive din kompetanse på kne- og hofteproteser?
- Har du kjennskap til informasjonssystemer som predikerer utfall av proteseoperasjoner? Dersom ja, kan du nevne disse?
- Er det vanlig å beregne hva som er årsaken til reoperasjoner for kne- og hofteproteser?
    - Hvordan gjøres dette?
    - Ser du på sammenhenger mellom proteseprodusenter og årsaker til reoperasjoner?

- Pleier du å gjøre vurderinger på kvaliteten på produktene til ulike proteseprodusenter?

- Hvordan gjøres dette?
    - Har du kjennskap til systemer som hjelper med slike vurderinger?

- Hvordan brukes databasene nå?
    - Hvordan ser du på dataen?
    - Hvordan syns du visualiseringen er i dag? Bra/dårlig?
    - Hjelper visualiseringene med forståelsen for dataen?
    - Er det kjente forbedringspunkt for visualiseringene?
- Hvordan syns du det fungerer?
    - Oppstår det noen problemer?
    - Hvis ja, hvordan løses disse?
- Hvilken funksjoner er tilgjengelig for deg nå?
    - Begrense søk, spesifisere søk, osv.
    - Er det noen funksjoner du savner?
- Hva er ditt hovedformål med registeret?
    - forskning, årsrapport eller daglig bruk?
- Hvor ofte kunne du tenkt deg å bruke registeret hvis dataen var tilgjengelig online?
    - Fra tid til annen/en gang i året?

- Utledende spørsmål; har du noe mer å tilføye som ikke har blitt spurt om?

Intervjuguide

Format: ansikt til ansikt eller digitalt intervju.

Svarregistrering: Lydopptaker, notater.

Hensikten med intervjuet er å først etablerte den faglige bakgrunn hos deltakerne, la deltakerne utforske og tolke resultatene og modellene jeg har utviklet, for deretter å få deltakernes oppfattelse av resultatene og hvorvidt de er aktuelle for fremtidige informasjonssystemer i ortopedi.

Intervju

Varighet rundt 30-60 minutter.

Deltakerne informeres om hva prosjektet går ut på og hva jeg ønsker med intervjuet. Deltakernes teknologiske kompetanse og deres kunnskap rundt artroplastikk og databaser kartlegges.

Spørsmål:

- Hvilken erfaring har du med teknologi?
    - Vil du si du er datakyndig (kompetent)?
- Hva er din høyeste utdanning?
- Hvilken holdning har du til teknologi generelt?
- Eier du en datamaskin/et nettbrett/en mobiltelefon?
    - Hva bruker du teknologier til på en daglig basis?
- Kan du fortelle om en vanlig dag for deg?
    - (hvor teknologi bli brukt)

- Hvordan vil du beskrive din kompetanse på kne- og hofteproteser?
- Har du kjennskap til informasjonssystemer som predikerer utfall av proteseoperasjoner? Dersom ja, hva er disse?

- Hvilken informasjon har du tilgang til rundt din protese?
    - Hvordan blir den presentert?
    - Hvilken informasjon er du mest interessert i?

- Har du tilgang til kne- og hofteregisteret?
- Hvis ja:
    - Hvilken data har du tilgang til?
    - Hvordan brukes databasene nå?
        - Hvordan ser du på dataen?
        - Hvordan syns du visualiseringen er i dag? Bra/dårlig?
        - Hjelper visualiseringene med forståelsen for dataen?
    - Hvordan syns du det fungerer?
        - Oppstår det noen problemer?
        - Hvis ja, hvordan løses disse?
    - Hvilken funksjoner er tilgjengelig for deg nå?
        - Begrense søk, spesifisere søk, osv.
        - Er det noen funksjoner du savner?
- Hvis nei:
    - Hvor ofte kunne du tenkt deg å bruke registeret hvis dataen var tilgjengelig online?
        - Fra tid til annen/en gang i året?

- Utledende spørsmål; har du noe mer å tilføye som ikke har blitt spurt om?

Vil du delta i forskningsprosjektet

"Maskinlæring i norsk register for kne- og hofteproteser"?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å lage en portal for analyse og datautvinning av data fra det norske artroplastikk registeret (NAR). I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål
Formålet med prosjektet er å utvikle en prototype av en portal hvor brukergrupper slik som ortopeder kan få tilgang til nyttig informasjon om kne og hofteoperasjoner. Hvis prototypen blir god nok, vil det bli utviklet et endelig produkt som tar utgangspunkt i prototypen.

Systemet vil bruke analyse og datautvinning teknikker på registerdata fra det norske artroplastikk registeret. Målet er å skape en plattform med innsikt og analyseredskap for å besvare kliniske spørsmål vedrørende pasienters hofte og/eller kneproteseoperasjon. Forskningsprosjektet vil ha to ulike, men relaterte og til dels overlappende fokus. 1. Et datautvinnings aspekt hvor det vil tas i bruk maskinlæring og andre dataanalyse teknikker for innsikt i data. 2. Et menneske-maskin-interaksjon (HCI) aspekt hvor fokuset ligger på brukergrensesnitt og hvordan en kan presentere analytisk data på en nyttig og intuitiv måte. Noen av spørsmålene vi ønsker å besvare er blant annet: Hva er den estimerte levetiden på protesen? Hvilke faktorer påvirker levetiden av en protese? Hvilke pasientgrupper assosieres med høyere grad av revisjonsoperasjoner? HCI relaterte problemstillinger inkluderer: Hvordan kan designmetoder fra HCI hjelpe med å visualisere data på best mulig måte for forskjellige brukergrupper? Hvordan kan sensitiv medisinsk data bli presentert på en etisk riktig måte? Kan bruk av HCI-metoder og -prinsipper på registerdata hjelpe leger, og forbedre pasientomsorgen?

Forskningsprosjektet består av et samarbeid mellom fire masterstudenter der hver student har ulike oppgaver tilknyttet prosjektet. Studentene har en felles veileder som skal resultere i fire individuelle masteroppgåver.

Hvem er ansvarlig for forskningsprosjektet?
Universitetet i Bergen er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Veileder har inngått et samarbeid på forhånd med forskere ved norsk register for leddproteser og ortopeder fra Haukeland Universitetssykehus. Utvalget er aktuelle brukere av det ovennevnte systemet. Som brukergruppe, kan de bidra med kravsetting og utforming av systemet. Som fagkyndig og praktiserende, kan de hjelpe med klinisk forståelse av behandlingen og generell rådføring gjennom forskningsprosjektet.

Hva innebærer det for deg å delta?

Hvis du velger å delta i prosjektet, innebærer det at du stiller til et personlig intervju. Det vil ta deg ca. 30-60 minutter. Intervjuet inneholder spørsmål om dine meninger eller oppfatninger om systemet. Vi ønsker tilbakemeldinger fra fagkyndige for å utvikle prototypen. Metoden vi vil bruke er personlig intervju. Det kan også bli aktuelt med brukertesting under observasjon.

Utvalg 1 (ortopeder), utvalg 2 (forskere), utvalg 3 (pasienter) og utvalg 4 (studenter) vil kunne bli bedt om å delta i et personlig intervju og brukertesting.

Det er frivillig å delta
Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

Det er ingen avhengighetsforhold for deg som deltaker.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger
Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- Prosjektgruppe og veileder.
- Navnet og kontaktopplysningene dine vil jeg lagre på en kryptert harddisk.

I utgangspunktet vil deltakere ikke kunne gjenkjennes i publikasjon, men kreditering av deltaker kan innfris etter ønske fra deltaker. I tilfelle det blir aktuelt, vil opplysninger som navn og yrke kunne bli gjengitt i publikasjon.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?
Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen skal skje juni 2020. Ved prosjektslutt vil alle personopplysninger bli anonymisert og alle typer opptak (lyd,video) vil bli slettet.

Dine rettigheter
Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg, og å få utlevert en kopi av opplysningene,
- å få rettet personopplysninger om deg,
- å få slettet personopplysninger om deg, og
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?
Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Universitetet i Bergen har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- Prosjektveileder Ankica Babic og/eller Arle Farsund Solheim (student) ved Universitetet i Bergen.
    - Ankica Babic - Kan nås på epost (Ankica.Babic@uib.no) eller på telefon: +47 55 58 91 39
    - Arle Farsund Solheim - Kan nås på epost (Arle.Solheim@student.uib.no)

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:

- NSD – Norsk senter for forskningsdata AS på epost (personverntjenester@nsd.no) eller på telefon: 55 58 21 17.

Med vennlig hilsen

Ankica Babic (Forsker/veileder)

Arle F. Solheim (Student)

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet *Maskinlæring i norsk register for kne- og hofteproteser*, og har fått anledning til å stille spørsmål. Jeg samtykker til:

- □ å delta i *personlig intervju*
- □ å delta i *brukertesting (observasjonsstudie)*
- □ at *Arle F. Solheim* kan gi opplysninger om meg til prosjektet
- □ at opplysninger om meg publiseres slik at jeg kan gjenkjennes gjennom *navn, yrke, kjønn, og alder.*

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet

----

(Signert av prosjektdeltaker, dato)

C Appendix - Evaluation forms

C.1 System Usability Scale

Participant ID: _____ Device: _____ Date: dd/mm/yy

System Usability Scale

Instructions: For each of the following statements, mark **one** box that best describes your reactions to the system *today*.

	Strongly disagree				Strongly agree
1. I think that I would like to use the system frequently					
2. I found the system unnecessarily complex					
3. I thought the system was easy to use					
4. I think I would need assistance to be able to use this system					
5. I found the various functions on the system well integrated					
6. I thought there was too much inconsistency on the system					
7. I would imagine that most people would learn to use the system very quickly					
8. I found the system very cumbersome/awkward to use					
9. I felt very confident using the system					
10. I need to learn to learn a lot of things before I could start using the system					

Please provide more comments about the system:

This questionnaire is based on the System Usability Scale, which was developed by John Brooke while working at Digital Equipment Corporation.

C.2 Nielsen's Heuristics

Nielsen's Heuristic

Date: dd/mm/yy

Device: ______

Severity:

0 – Everything is fine

1 – Cosmetic problem only: need not be fixed unless extra time is available

2 – Minor usability problem: fixing this should be given low priority

3 – Major usability problem: important to fix, should be given high priority

4 – Usability catastrophe: imperative to fix this before product can be released

Heuristic	Severity	Issues	Recommendation
Visibility of system status The system should always keep users informed about what is going on, through appropriate feedback within reasonable time			
Match between system and the real world The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.			
User control and freedom Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.			
Consistency and standards Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.			
Error prevention Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.			
Recognition rather than recall Minimize the user's memory load by making objects, actions, and options			

C.3 Content Evaluation Table

Content of the Data Mining, Visualization, and HCI Functionality

	Functionality	**Totally disagree**	**Disagree**	**Neutral opinion**	**Agree somewhat**	**Totally Agree**
1	Choice of Data Mining (DM) tasks					
2	Need to add additional tasks					
3	Welcoming starting page with something like demographics					
4	Save all DM sessions					
5	Choice of visualization					
6	Level of interactivity in visualization					
7	HCI outlay is satisfying					
8	Need to add additional HCI functionality					
9	HCI interface is well suited for experts					
10	HCI interface has potential to meet patient needs					

Add your comments below following the numbers:

D Appendix - Licences

D.1 Codepen Licence

D.2 SVG Map Library License

SVG Map Library License | Resources

You are free to use our SVG maps, without any restrictions in web sites, web templates and software programs of any kind that you intend to sell or distribute. Backlinks are not required, but are highly appreciated. **You are free to use our SVG maps for personal or commercial use.** The only caveats are as follows:

You may not make our SVG maps "as is" available for distribution elsewhere without prior permission. In other words, **you can't copy a collection of our maps, not add any value, and distribute them as your own.** If you'd like to feature our map collection, please do so by linking directly to: https://simplemaps.com/resources/svg-maps.

This library is provided "as-is". We make no guarantees as to the reliability of the data or its fitness for any particular use. By using this library, you agree not to hold us liable for any damages that may arise.