Developing Conceptual Designs for Digital Twins in Arthroplasty

Ida Wergeland Sævareid

Ankica Babic

A Master's Thesis
Department of Information Science and Media Studies
University of Bergen

June 3, 2024

Abstract

This research has been dedicated to finding visualizations for the novel and rather abstract concept of the digital twin, which is often presented either as a flashy color figure or as some kind of model or results from data analysis.
The idea of a twin is to summarize what defines a group of subjects based on their nearly identical features. In the health sector, we often consider patients and what constitutes good or optimal treatment for expected outcomes. We define different patient groups and also try to connect them with some form of decision support. In the case of the digital twin, the project aimed to illustrate how to identify a digital twin for a specific patient and assess the likelihood of a beneficial outcome.
In the field of arthroplasty, an attempt was made to create a digital twin by applying cluster analysis, event log analysis, and other methods to understand the potential clinical pathway a patient can take.
This project, through five iterations, generated several conceptual designs of digital twins. One design related a patient to a cluster where a digital twin could be found based on the similarity of the data. Another aimed to find a path that a patient is likely to follow based on their baseline data. The third model integrates both clusters and pathways to provide a more holistic picture of where a digital twin and a likely clinical pathway could be related.
A digital twin is a novel concept and its visualization relies on data analytical methods to extract the most significant features. The conceptual solutions of this project are general and allow for the inclusion of different kinds of results while maintaining a predefined way of navigating and connecting data. This helps users retain a consistent understanding of creating digital twins and interpreting the results. Cognitive walk-throughs were applied to assess the usability of these conceptual designs, which are in their final iteration as mid-fidelity prototypes.
Working with two different datasets, a quality data register and an intensive care unit dataset, has shown differences in the structure and intensity of information. The challenge is to create conceptual designs that can suit both datasets, but one feasible solution is to create different instantiations for the same conceptual design. Future work involves developing a fully functioning high-fidelity prototype that can operate on different databases while providing the same sense of the digital twin concept.

Acknowledgements

I would like to thank William and Carl for excellent collaboration on this project, from sharing data to providing me with feedback.

A special thank you goes to Dr. Ankica Babic, University of Bergen, for her guidance and support throughout the project.

I also wish to thank Dr. Peter Ellison from the University of York for his insights into the NAR database and for his feedback during the early stages of the design.

Heartfelt thanks to my parents for always supporting me, for reading my thesis, and for ensuring that I ate well.

Last but not least, thanks to my roommates at 642 for being there through it all, sharing both frustrations and laughter.

Ida Wergeland Sævareid
Bergen, 31/05/2024

Contents

List of Figures

List of Tables

Abbreviations and Definitions

AI: Artificial Intelligence.

ASA score: The American Society of Anesthesiologists.
A classification system that determines the health of a person prior to a surgical procedure that requires anesthesia.

BMI: Body Mass Index.
Persons weight in kilograms divided by the square of heights in meters.

CCI: Charlson Comorbidity Index.
Predicts the mortality for a patient who may have a range of concurrent conditions.

EHR: Electronic Health Records.

HIPAA: Health Insurance Portability and Accountability Act.

Industry 4.0: Fourth Industrial Revolution.
Phrase describing rapid technological advancement in the 21st century.

ICD-code: International Classification of Diseases.

ICU: Intensive Care Unit.

IoT: Internet of Things.

MIMIC-IV: Medical Information Mart for Intensive Care IV.

MIT: Massachusetts Institute of Technology.

NAR: Norwegian Arthroplasty Registry.

Prosthesis survival: *A number of years that patient has a prosthesis.*

RQ: Research Question.

Chapter 1

Introduction

1.1 Introduction

Advancements in technology have revolutionized various industries, and the healthcare sector is no exception. The concept of digital twins, originally developed in engineering, has emerged as a promising tool to understand and simulate complex systems in healthcare. Digital twins refer to virtual representations of real-world objects, processes, or individuals, allowing for comprehensive analysis, monitoring, and prediction in a digital environment [3]. This Master thesis explores the application of digital twins in healthcare, specifically focusing on visualizing patient data obtained from the Norwegian Arthroplasty Register (NAR) and Medical Information Mart for Intensive Care IV (MIMIC-IV).

NAR is a comprehensive database that collects valuable data on joint replacement surgeries in Norway [14], while MIMIC-IV provides deidentified intensive care unit (ICU) data from patients at Beth Israel Medical Center [18]. Leveraging these extensive datasets, the main task of this thesis is to develop and explore the possibilities of visualizing patients' characteristics, their digital twins, and the clinical pathway of the different digital twins. Using design science as the primary methodological approach, conceptual designs and dashboard prototypes have been created to facilitate the analysis and use of patient data alongside their corresponding digital twins.

The primary objective of this research is to contribute to understanding and improving patient outcomes in joint replacement surgeries through the implementation of digital twins. The use of digital twins enables healthcare professionals to gain insight into patients' preoperative conditions, understand the impact of various factors, and potentially predict surgical outcomes. By visualizing these data-driven digital twins alongside patient-specific information, a comprehensive dashboard prototype has been developed to support clinicians in their decision-making process and provide personalized treatment pathways.

This study will encompass various elements including data analysis, design, and development of the digital twin framework, and the implementation of a dashboard prototype. The findings of this research are expected to not only improve healthcare outcomes for hip replacement surgery patients, but also contribute to the growing body of knowledge within the field of digital health and inform future advancements in leveraging digital twins in healthcare applications.

Overall, this thesis aims to bridge the gap between data-driven decision making and clinical practice by integrating digital twins into healthcare. Through visualization of patient data from the NAR and MIMIC-IV databases, the proposed conceptual design and dashboard prototype will improve clinical understanding, facilitate evidence-based decision making, and potentially improve patient outcomes.

1.1.1 Collaboration

This research is conducted in collaboration with two other master students. Sahlgaard [30] and Røise [29] have worked on the back-end part of the project, performing various data mining and data analyzes to create results. These results have been central in the visualization and front-end part presented in this thesis.

1.1.2 Problem statement

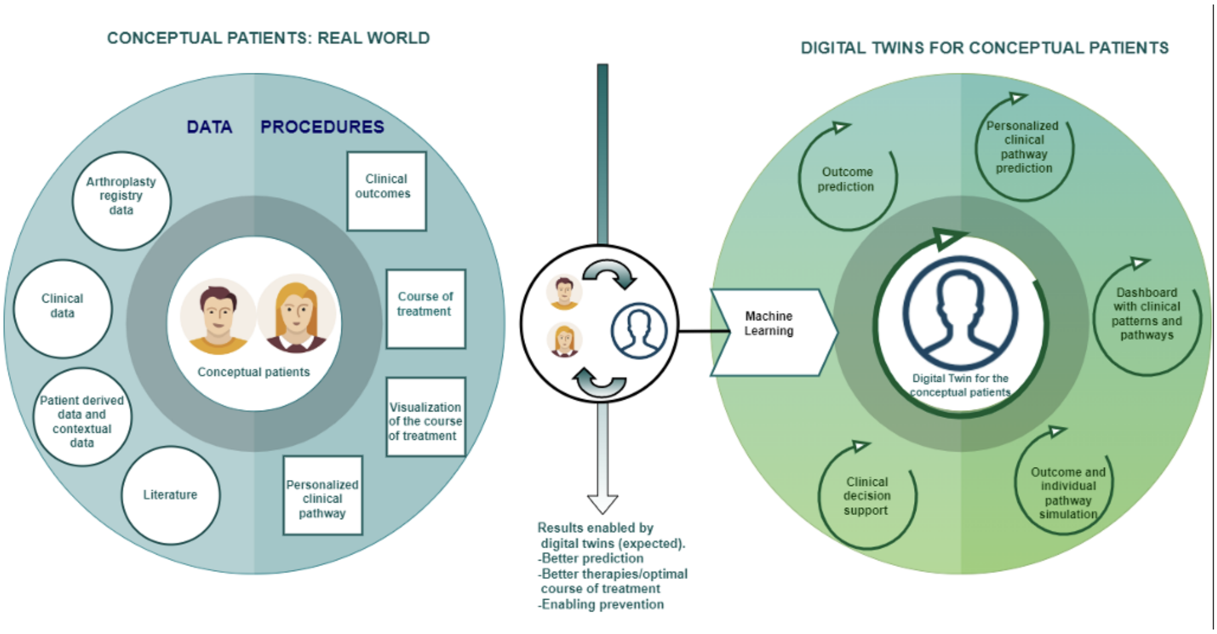

Figure 1.1: The Digital Twin project [4].

Figure 1.1 presents the framework for the digital twin project. On the left, it illustrates the real-world clinical data and procedures that physicians use to understand patient outcomes, treatment courses, and individualized patient pathways. This data includes information from the arthroplasty registry, patient records, patient-generated data, and direct patient data collection.
On the right, the framework shows several modules created using machine learning methods. These modules are essential for developing models that support physicians in managing treatment courses, defining individualized pathways, and visualizing critical steps. These models are referred to as digital twins because they represent real-world patients through their data and the relationships between the data points.
The visualization task in this project operates on multiple levels. It includes schematic representations that cover the data, operations, and models involved in creating digital

twins. Additionally, various visualization methods are employed to provide a comprehensive understanding. My research will focus on outcome analysis and processes, complementing the work done by Sahlgaard [30] and Røise [29].

1.1.3 Research questions

My research question is divided into three parts. The first two questions are connected to the work of the two other students on the project. The third question concerns the issues addressed in the first two questions. Furthermore, it should be noted that the term "arthroplasty" in this context refers to the datasets.

1. **What is the best way to visualize the arthroplasty outcome analysis in the Digital Twin concept?**
   - This question focuses on the visualization connected to the outcome of the patients, and needs to show a current picture of the digital twins.
2. **What is the best way to visualize the arthroplasty clinical pathway of the Digital Twin concept?**
   - The second research question deals with the pathway of a patient, and the details around this aspect of the data.
3. **What are the alternatives for presenting the whole arthroplasty Digital Twin combining various analysis?**
   - The last question deals with the whole visualization part, and focuses on how to best gather the different information from the first two questions. It focuses on the overarching visualization aspect; concentrating on the optimal integration of dashboards.

To address these research questions, I have explored various visualization methods. Subsequently, I have engaged in a design process that culminates in the development of conceptual designs and a prototype, which is then evaluated and tested in the context of the research questions.

1.2 Motivation

There is no (known) detailed visualization on digital twins based on data analysis. The current literature seems to offer an overview presentation in which a digital twin is presented symbolically or with flashy figures (neon-alike). My intention has been to explore how to get closer to the digital twin using the results of real data and analyses based on the data.

Chapter 2

Background

A digital twin is a virtual replica of a real-world item, individual, or process, situated within a virtual representation of its environment. This concept, fundamental to Industry 4.0, allows organizations to model real-world scenarios and predict outcomes, thus enhancing decision-making processes [35]. Using simulation, machine learning, and reasoning, a digital twin serves as a virtual counterpart that spans the entire lifecycle of an object or system, continuously updated with real-time data [3].

In the field of medicine and healthcare, digital twins are particularly valuable as they need to represent the current state of a patient and predict future course of action [1]. This requires clear and comprehensible visualizations. For this project, the objective is to explore how patient data, existing and new, can be used with this technology.

The digital twin is a novel concept and without context rather abstract. The literature reveals various attempts to define this concept through data interpretation and analysis, leading to work definitions of digital twins that are often context-specific. This makes visualization challenging, as many authors present models and flashy figures to represent the concept of a digital twin.

In this work, a digital twin is presented in the clinical context of arthroplasty and analytically through data mining and the creation of clinical pathways. The resulting conceptual designs illustrate how to derive a digital twin based on current patient data. For instance, patients can be categorized into different clusters, each representing potential pathways leading to various outcomes, ranging from successful recovery to complications related to surgery.

These conceptual designs provide a visual representation of different scenarios, allowing physicians or researchers to interpret potential outcomes based on patient data. The designs help make the patient's outcome more transparent (through clustering) and sequential (through pathways). My thesis work builds on the concrete analyzes conducted by other team members, Sahlgaard [30] and Røise [29], which significantly influence the conceptual designs.

Thus, my definition and understanding of a digital twin are tailored to this specific arthroplasty application, which is most relevant to medical professionals and researchers in the field. Although my research focuses on arthroplasty, the conceptual design may also be applicable to similar research challenges involving data mining and the digital twin concept.

2.1 Literature overview

In this section, the relevant literature is presented. The articles are categorized into Digital Cousins, Visualizations of Digital Twins, Digital Twins in Healthcare, Dashboard Design, and Visualization of Clinical Pathways. But first a definition of arthroplasty is given.

2.1.1 Arthroplasty

Arthroplasty is a surgical procedure that restores the function of a joint. A joint can be restored by resurfacing the bones or using an artificial joint (prosthesis) [22].
Hip arthroplasty, also known as hip replacement surgery, is a common procedure in which a damaged hip joint is replaced with a prosthetic implant. This surgery is often recommended for individuals suffering from severe hip pain and mobility problems due to conditions such as arthritis, fractures, or other degenerative joint diseases. The procedure involves removing damaged parts of the hip joint and replacing them with prosthetic components designed to mimic the function of a natural hip joint [23].

2.1.2 Digital Cousins

Cousins, Siblings and Twins: A Review of the Geological Model's Place in the Digital Mine [16]

The article discusses the challenges and impact that the environment has on digital twins, particularly in the context of natural systems such as mining environments. Although engineered systems can be digitally twinned, natural systems pose unique uncertainties.
The minimum criterion for a digital twin is a synchronized, real-time pairing of virtual and physical domain capable of predicting behavior and informing decision-makers for productivity and safety. Digital twins range from representing single parts of a system to entire machines, with a two-way link between physical and virtual systems allowing for adjustments and informed decision-making.
A digital cousin, on the other hand, maintains a distant relationship with the physical system. It relies more on qualitative and interpreted data with expert judgement and less on quantitative data. Another concept presented is the digital sibling. Compared to a digital cousin, a digital sibling is more closely related to the physical system since it contains more quantitative data instead on some qualitative data, but unlike the twin, it still has estimates and judgements.

2.1.3 Visualizations of Digital Twins

Integrated knowledge visualization and the enterprise digital twin system for supporting strategic management decision [38]

Yan, Hong, and Warren [38] discuss the application of integrated knowledge visualization and the concept of a digital twin of the enterprise to facilitate strategic decision

making within organizations. The authors emphasize the importance of effective decision making in today's complex and rapidly changing business environment.
The concept of a digital twin refers to a virtual representation of a physical object or system. In the case of an enterprise digital twin, it encompasses the entire organization, including its processes, people, products, and services. The authors argue that the integration of knowledge visualization techniques with an enterprise digital twin can provide decision makers with a comprehensive view of the organization's current state and future scenarios.
The article highlights several key benefits of the proposed approach. First, it enables decision makers to obtain dynamic and real-time information about various aspects of the organization. This allows for a better understanding of the complex interdependencies between different parts of the enterprise and the impact of potential decisions.
Secondly, the integration of knowledge visualization techniques improves decision-making by employing visual methods such as infographics, charts, and graphs. These visuals can convey complex information in a more understandable and concise manner, allowing decision makers to recognize patterns, trends, and relationships that might not be apparent in traditional textual formats.
Furthermore, they discuss the use of data analytics and simulation within the enterprise digital twin system. By analyzing large volumes of data and running simulations, decision-makers can assess the potential consequences of different strategic decisions. This helps identify risks, evaluate alternatives, and make informed decisions that align with organizational goals.
The authors provide a case study that illustrates the practical implementation and benefits of the integrated knowledge visualization and the enterprise digital twin approach. The case study involved a multinational manufacturing company, where the system facilitated the identification of production bottlenecks and the optimization of resource allocation, ultimately resulting in increased efficiency and profitability.
In summary, the article highlights the potential of combining knowledge visualization techniques with a digital twin in an enterprise to support strategic decision making. The integration of real-time information, visual methods, and data analytics enables decision makers to gain a comprehensive understanding of their organization and make informed decisions that can lead to improved performance and competitive advantage.

2.1.4 Digital Twins in Healthcare

Digital twins to personalize medicine [7]

The article discusses the potential of using digital twin technology to revolutionize healthcare and personalize medical treatments. The authors refer to a digital twin as a virtual replica of a physical entity or system created using real-time data and advanced analytics. In the context of personalized medicine, digital twins can be created to represent an individual patient, incorporating their unique biological characteristics, medical history, and genetic information.
It highlights that by using digital twin technology, healthcare professionals gain a better understanding of each patient's unique characteristics, allowing for more precise diagnosis and treatment plans. Rather than using generalized approaches to medical care, digital twins enable personalized strategies customized to the specific needs of each in-

dividual. This is echoed by Angulo et al. [1], who argue that digital twins enable a comprehensive understanding of patients and offer holistic views of their health.

One of the primary benefits of digital twins in medicine is their ability to simulate the effects of different treatments on the virtual representation of a patient. By monitoring and adjusting the variables of the digital twin, such as drug dosage, duration of treatment, or lifestyle changes, physicians can evaluate potential outcomes and refine treatment plans before applying them to the actual patient. This proactive approach to treatment is further supported by Shengli et al. [32], who discuss the feasibility and benefits of human digital twins in identifying health risks and recommending preventive measures.

Moreover, digital twins can also help predict disease progression and assess the effectiveness of preventive measures. By continuously monitoring the digital twin and feeding it with real-time data, healthcare professionals can anticipate potential health deteriorations and take proactive steps to prevent them. Angulo et al. [1] emphasize the potential of digital twins in early detection and disease prevention through continuous monitoring and pattern recognition.

"Digital twins to personalize medicine [7] further explains that digital twins have the potential to streamline medical research and innovation. Researchers can use digital twins to simulate the effects of experimental drugs or therapies, reducing the reliance on animal testing and accelerating the drug discovery and development process. This point is reinforced by Armeni et al. [2], who highlight the potential of digital twins to accelerate drug responses and improve treatment efficacy through patient-specific simulations.

In addition, digital twins have the potential to empower patients with insights about their own health and well-being. By granting individuals access to their digital twin data, patients can make informed decisions about their lifestyle choices, monitor their own health, and actively participate in their treatment plans. However, the article also highlights some challenges and concerns associated with the widespread implementation of digital twins in healthcare. These include the need for robust data security and privacy measures, extensive computational power and storage, and ethical implications of controlling and manipulating digital representations of human lives. These challenges are further elaborated by Sun et al. [35], who stress the importance of maintaining data integrity and security to gain patient trust.

In conclusion, the article sheds light on the promising potential of using digital twin technology in healthcare. By leveraging these virtual replicas, personalized medicine can reach unprecedented levels, improving patient care, treatment outcomes, and medical research. Nonetheless, careful consideration must be taken to address the challenges and ethical concerns to ensure the responsible and ethical application of digital twin technology in healthcare.

On the Integration of Agents and Digital Twins in Healthcare [11]

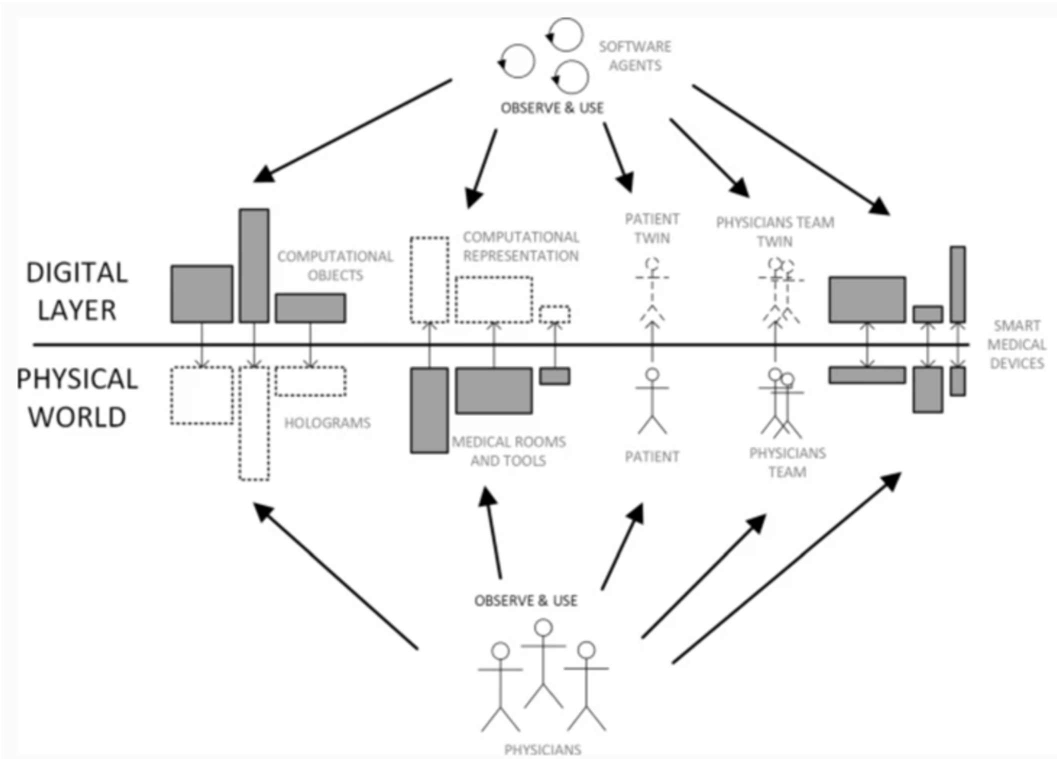

Figure 2.1: Agent-based digital twins as mirror worlds - a conceptual representation for the healthcare context [11].

Croatti et al. [11] explores the potential benefits and challenges of integrating agents and digital twins in the healthcare domain. The purpose of this project is to provide a comprehensive understanding of these technologies and their implications in improving patient care and healthcare delivery.
Agents refer to intelligent software modules that can autonomously perform specific tasks or make decisions. They have the ability to communicate, collaborate and negotiate with other agents to achieve their goals. Digital twins are virtual replicas of physical entities or systems, created using real-time data and simulations. In healthcare, digital twins can represent individual patients, organs, or entire healthcare settings.
The integration of agents and digital twins in healthcare opens up new possibilities for personalized and proactive care, as it enables the real-time monitoring, analysis, and prediction of patients' conditions. Using digital twins as the foundation, agents can be deployed to collect data, analyze it, and make informed decisions. For instance, an agent can monitor patient vital signs through various sensors connected to the digital twin of the patient's body, and alert healthcare providers in case of anomalies or critical situations.
Furthermore, agents and digital twins have the potential to optimize resource utilization in healthcare settings. Agents can collaboratively manage the allocation of healthcare resources, such as hospital beds, medical equipment, and staff, based on real-time information provided by digital twins. This can lead to better operational efficiency, reduced waiting times, and improved patient outcomes. Sun et al. [34] highlight similar benefits, emphasizing the role of digital twins in optimizing treatment plans and improving

patient outcomes through predictive modeling.
However, the integration of agents and digital twins in healthcare also poses challenges. One major concern is the privacy and security of patient data. As digital twins accumulate sensitive information about patients, adequate measures must be in place to ensure data protection and prevent unauthorized access. Another challenge is the complexity of implementing such integration in existing healthcare infrastructure, which often includes legacy systems and interoperability issues. Sun et al. [34] also discuss the integration challenges, particularly focusing on the need for standardization and interoperability between various healthcare systems and digital twin models.
In conclusion, the article highlights the potential benefits of integrating agents and digital twins in healthcare, including personalized care, proactive monitoring, and optimized resource allocation. However, it also emphasizes the need to address challenges related to privacy, security, and system integration. Overall, the integration of these technologies has the potential to revolutionize healthcare delivery and significantly improve patient outcomes.

Digital twin in healthcare: Recent updates and challenges [34]

Sun, He, and Li [34] discuss the concept of digital twins in the healthcare sector, providing an explanation and summarizing the recent advancements and obstacles associated with its implementation.
They begin by introducing the concept of digital twins, which essentially comprises a virtual replica of a physical object or system. When applied to healthcare, digital twins play a pivotal role in offering personalized and precise care to patients. By combining patient-specific data, such as medical records, genetics, lifestyle, and behavioral information, with sophisticated algorithms and predictive models, digital twins can simulate and predict the impact of various interventions and treatments on individual patients.
The article highlights recent updates in the field, emphasizing the increasing utilization of digital twins in disease prevention, diagnosis, treatment optimization, and post-treatment monitoring. It mentions the successful application of digital twins in areas such as cardiovascular health, orthopedics, neurology, and even mental health. The ability of digital twins to simulate the effects of medications, medical devices, and surgical procedures before their implementation enables healthcare providers to make informed decisions and improve patient outcomes. Armeni et al. [2] support this view, noting that digital twins can accelerate drug responses and enhance treatment efficacy through patient-specific modeling.
However, the article sheds light on the challenges associated with the widespread adoption of digital twins in healthcare. Privacy and security concerns about the vast amount of personal data required to create accurate digital twin models are a major obstacle. Maintaining data integrity, ensuring secure transmission and storage, and obtaining patient consent for data usage are crucial aspects that need to be addressed to gain patient trust. The integration of digital twin technology into existing healthcare systems is another challenge mentioned. The seamless interoperability of digital twin models with Electronic Health Records (EHR) and other clinical systems is essential for their practical implementation. Standardization of data formats, protocols, and interfaces between various healthcare stakeholders is necessary to achieve this integration. Angulo et al. [1] and Shengli et al. [32] also emphasize the importance of addressing privacy, secu-

rity, and integration challenges to ensure the responsible deployment of digital twins. Furthermore, the challenge of handling the vast data complexity associated with a digital twin is discussed. The integration of numerous data sources, including wearable devices, sensors, genomic sequencing, and patient-reported results, poses difficulties in data processing, analysis, and scalability.
In conclusion, the article provides an overview of the concept of digital twins in healthcare and highlights its recent advancements and obstacles. Despite the challenges of privacy, integration, and data complexity, digital twins hold great promise for personalized and effective healthcare. With further advances in technology and collaborative efforts across the healthcare ecosystem, digital twins have the potential to revolutionize patient care and outcomes.

2.1.5 Dashboard Design

Developing a Data Dashboard Framework for Population Health Surveillance: Widening Access to Clinical Trial Findings [10]

Concannon et al. [10] discuss the development and evaluation of a data dashboard framework for population health surveillance. The framework was developed to visualize datasets produced at a demographic health surveillance site in South Africa. The aim was to create a comprehensive, reusable, and scalable dashboard design framework that would provide increased access to information and facilitate evidence-based decision making.
The framework consisted of five core considerations: study setting, purpose and concept of the dashboard, user interaction and flow, design of data selection and visualization, and framework architecture. The study setting was the Africa Health Research Institute (AHRI), which operates as a demographic surveillance site. The purpose of the dashboard was to provide users with access to information on the studies on the site and to increase the visibility and accessibility of the datasets produced. The dashboard allowed users to monitor key performance indicators, explore datasets spatially, and compare global and local performance.
When creating a dashboard, especially in the context of healthcare, there are several main things to consider.

1. *Study setting:* Understand the specific context in which the dashboard will be used, including the type of data being collected and the privacy and confidentiality concerns.

2. *Dashboard purpose and concept:* Determine the purpose of the dashboard, the target user group, and the objectives of the dashboard. Consider what information needs to be displayed and how users will interact with the dashboard.

3. *User interaction and flow:* Define the intended process of user interaction with the dashboard interface. Consider how users will navigate through the dashboard and access the information they need.

4. *Data selection and visualization design:* Select the key indicators that will be displayed on the dashboard and consider the best ways to visualize and represent

these indicators. Choose visualization methods that are clear and intuitive to the users.

5. *Framework architecture:* Consider the technical aspects of the dashboard framework, including the input data structures and interfaces. Ensure that the dashboard is designed in a way that allows for easy updating and future scalability.

Overall, it is important to design a dashboard that meets the information needs of the specific healthcare context, while also being user-friendly and easy to navigate. In addition, considerations should be made for the privacy and security of the data being displayed.

Designing Attentive Information Dashboards [36]

Toreini et al. [36] focuses on the importance of designing effective information dashboards that capture user attention. It emphasizes the necessity for designers to create dashboards that not only present relevant information but also engage users and promote efficient decision-making.
One main argument discussed is the need for attention-driven design. They argue that traditional dashboards often overload users with excessive information, leading to information overload and decreased user engagement. They propose a shift towards attentive design, which involves prioritizing and presenting information in a manner that grabs users' limited attention spans and ensuring that the most critical information is highlighted effectively.
Another argument put forth revolves around the concept of dashboard aesthetics. The authors contend that aesthetics play a significant role in capturing user attention and improving overall user experience. They argue that visually pleasing dashboards with well-designed layouts, color schemes, and typography can enhance user engagement and make information more easily accessible. This argument emphasizes the importance of considering the visual appeal of a dashboard alongside its functional aspects.
Furthermore, the article discusses the importance of providing contextual information within dashboards. The authors argue that simply presenting raw data without any meaningful context can hinder users' understanding and decision-making processes. They suggested incorporating visual cues, annotations, and contextual information to help users gain a more holistic understanding of the data presented. This argument highlights the need to go beyond static data visualization and provide users with relevant background information to facilitate interpretation.
Lastly, the article stresses the value of user-centered design in the creation of information dashboards. The authors argue that designers should invest in understanding users' goals, needs, and preferences to tailor dashboards accordingly. They propose involving users in the design process through iterative feedback loops to ensure that the dashboards effectively meet their requirements. This argument emphasizes the importance of user research and usability testing to create dashboards that are intuitive, useful, and engaging.
In summary, the main arguments of the article revolve around the following points: the need for attention-driven design, the significance of aesthetics in capturing users' attention, the importance of providing contextual information, and the value of user-centered design. Understanding and implementing these arguments can help designers

create more effective information dashboards that engage users and facilitate optimal decision-making.

Dashboard Design Patterns [5]

Bach et al. [5] discuss various key design principles and patterns to create effective and visually appealing dashboards. It emphasizes the importance of designing dashboards that provide clear and actionable information to users in a concise and visually appealing manner.
It begins by highlighting the purpose of dashboards, which is to present data and information in a way that enables users to make informed business decisions. It emphasizes that an effective dashboard design should focus on presenting relevant and meaningful data, rather than overwhelming users with excessive information.
An important design pattern discussed in the article is the concept of information hierarchy. It suggests organizing the elements of the dashboard in a hierarchical manner, with the most important high-level information prominently presented at the top. This helps users quickly grasp the overall picture before delving into more specific details.
Another key principle highlighted is the use of visual cues and data visualization techniques. The article emphasizes the importance of using charts, graphs, and other visual elements to present data in a meaningful and easily understandable way. It also discusses the importance of color schemes, fonts, and other visual design elements in creating a user-friendly and visually appealing dashboard.
Another element presented is the need for consistency in dashboard design. It recommends using consistent design patterns, color schemes, and layouts throughout the dashboard to create a cohesive and intuitive user experience. Additionally, proper use of white space and decluttering the dashboard by removing unnecessary elements are suggested to improve usability.
Furthermore, the authors touch on the importance of responsiveness and adaptability in dashboard design. With the increasing use of mobile devices, it is crucial to create dashboards that can adjust and scale accordingly across different screen sizes and resolutions.
In conclusion, the article provides a comprehensive overview of key principles and patterns for creating effective dashboards. By focusing on information hierarchy, visual cues, consistency, and responsiveness, designers can create dashboards that effectively communicate data and facilitate decision making for users.

A Framework for Evaluating Dashboards in Healthcare [39]

The article discusses the importance of dashboards in the healthcare industry and proposes a framework for evaluating these dashboards. The authors highlight that dashboards play a crucial role in supporting decision-making processes in healthcare, as they provide visual representations of data for healthcare professionals.
It emphasizes that an effective dashboard should not only present data, but also assist in understanding and interpreting the information. To achieve this, the authors propose a framework consisting of six dimensions to evaluate healthcare dashboards. These dimensions include data quality, visual representation, functionality, user experience, data integration, and impact on decision-making.

The first dimension, *data quality*, emphasizes the accuracy, completeness, and consistency of the data presented on the dashboard. It highlights the importance of reliable data to make informed decisions.
The second dimension, *visual representation*, focuses on how the data is displayed on the dashboard. It suggests that the dashboard should utilize intuitive and visually appealing designs to enhance user understanding and engagement.
The third dimension, *functionality*, addresses the features and capabilities of the dashboard. It suggests that the dashboard should allow users to interact with the data, customize views, and access additional information.
The fourth dimension, *user experience*, evaluates the overall usability and satisfaction of the dashboard. It emphasizes the importance of an intuitive interface that enables users to navigate, search, and retrieve the desired information easily.
The fifth dimension, *data integration*, highlights the importance of integrating data from multiple sources into the dashboard. It emphasizes that the dashboard should provide a comprehensive view of the patient's health information, incorporating data from various healthcare systems.
Lastly, the sixth dimension, *impact on decision making*, examines the influence of the dashboard on the decision-making processes of healthcare professionals. It suggests that an effective dashboard should positively impact decision-making outcomes and promote evidence-based practices.
By assessing healthcare dashboards based on these six dimensions, the proposed framework aims to provide a comprehensive evaluation process. It helps identify strengths and weaknesses in existing dashboards and guides the design and development of improved dashboard solutions.
In conclusion, the framework proposed in the article addresses the need for an effective evaluation of healthcare dashboards. It emphasizes the importance of data quality, visual representation, functionality, user experience, data integration, and impact on decision making. By promoting the use of a well-designed dashboard, healthcare professionals can make more informed decisions, improve patient care, and improve overall healthcare outcomes.

2.1.6 Visualization of Clinical Pathways

Visualization of key factor relation in clinical pathway [37]

Yamashita et al. [37] discuss a method for visually representing the relationships between key factors in clinical pathways. Using a network-based approach, the authors developed a system that can help healthcare professionals understand the complex interactions between factors such as clinical guidelines, treatment options, and patient outcomes. The system allows for a more intuitive understanding of these relationships, which can ultimately lead to improved patient care and outcomes.

Chapter 3

Datasets

In the digital twin project, we have used two different dataset; Norwegian Arthroplasty Registry (NAR) and the Medical Information Mart for Intensive Care IV (MIMIC-IV). Both databases include data about arthroplasty patients.

3.1 The NAR database

The NAR database collects information from all hospitals in Norway on joint arthroplasties dating back to 1987 [14]. Its primary objective is to ensure that all patients receive optimal treatment. The database includes data about 10.000 hip arthroplasty patients. By analyzing data from the registry, the objective is to evaluate the efficacy of various prostheses and surgical techniques used throughout Norway. The ultimate goal is to quickly detect and discontinue the use of poorly performed implants and procedures, thereby improving patient outcomes.
In the digital twin project, we have explored the possibility of leveraging this extensive dataset to create digital twins for predictive modeling and personalized healthcare interventions. Specifically, my role has involved creating understandable visualizations to not only comprehend the dataset itself, but also interpret the results from Sahlgaard [30] and Røise [29].

Initial overview

As a part of the get-to-know-the-dataset phase, I have created some graphs and visualizations to get an overview of the demographics within the NAR database. In the following, the gender, age, county and survival of the prosthesis will be presented.

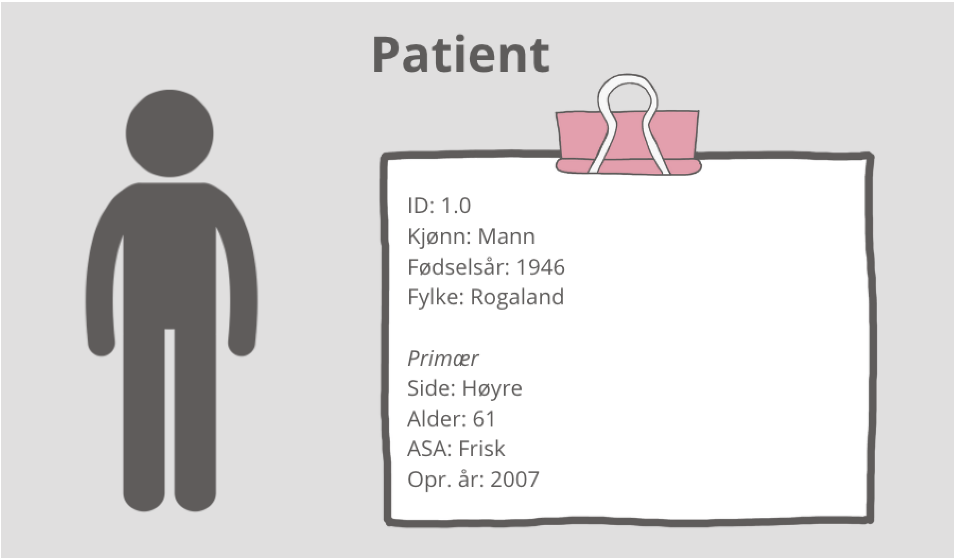

Figure 3.1: Primary patient status in brief

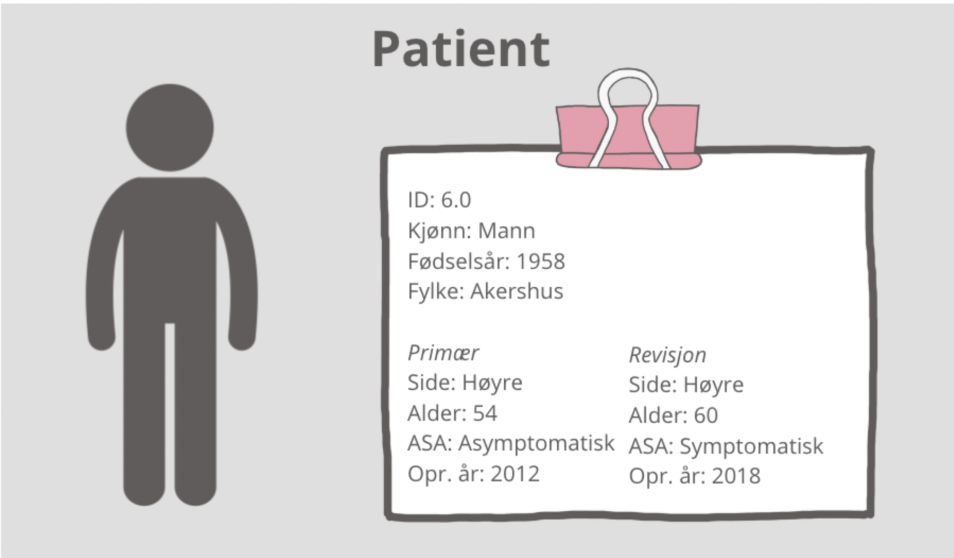

Figure 3.2: Revision patient status in brief

Figures 3.1 and 3.2 are patient status of two cases in the dataset, to represent the available information. The first patient has a desired outcome, while the second reoperated after only six years.

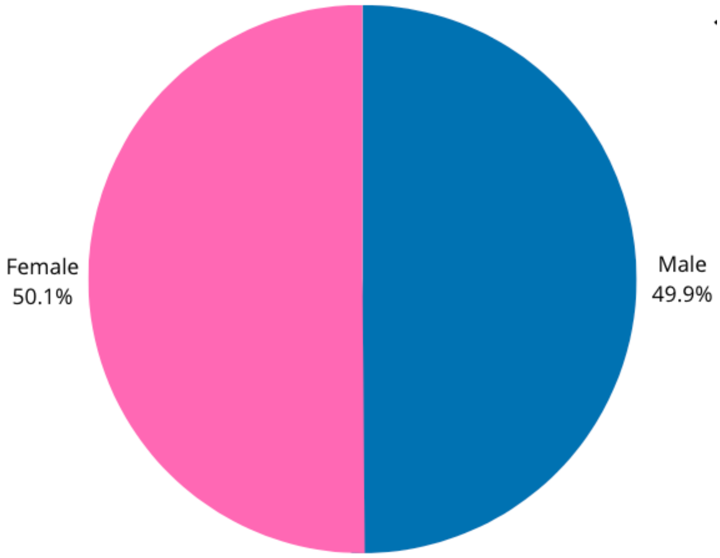

Figure 3.3: Gender distribution in the NAR database

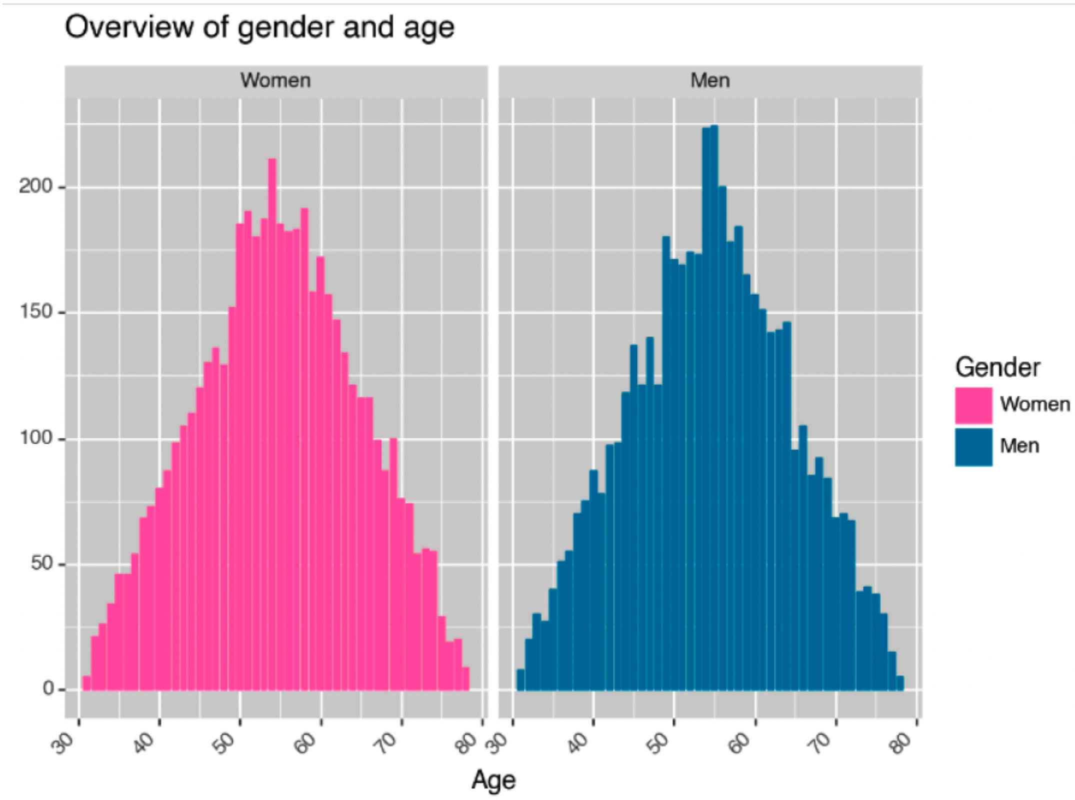

Figure 3.4: Gender and age distribution in the NAR database

Figures 3.3 and 3.4 are general descriptions of gender and age within the NAR database. It shows an almost identical separation in gender. The age distribution is centered around 50-60 years, with a good number of cases throughout the ages as well.

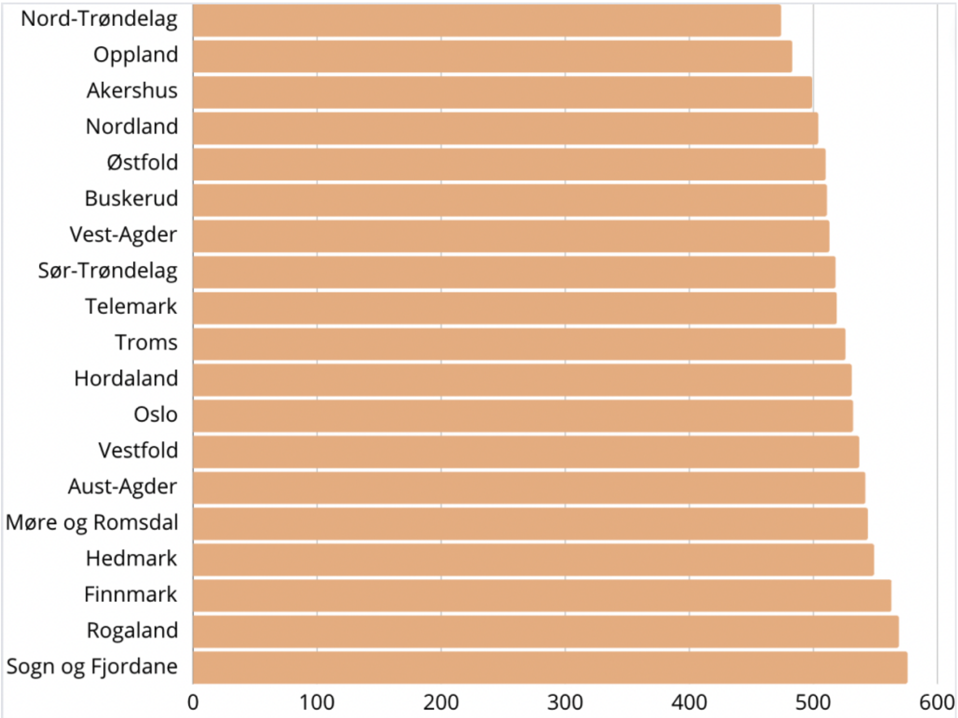

Figure 3.5: County distribution in the NAR database

Figure 3.5 shows an overview of from which county the patients in the database have done their operation. Ones again is the distribution very evenly separated and the operations are divided well throughout Norway.

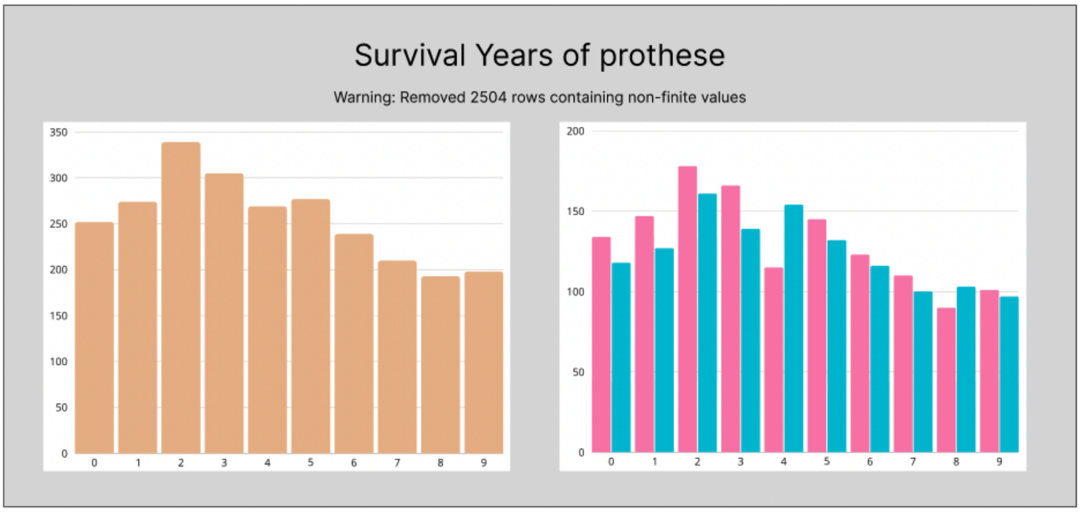

Figure 3.6: Prosthesis survival distribution in the NAR database

Figure 3.6 shows the distribution of how long survival-time the prostheses in the database have. Prosthesis survival of overall population on the left and divided by gender on the right. Overall, women are of higher representation in the section for 0-3 years survival. In addition, note that 2504 rows have been removed from the overview as these had non-finite values.

3.2 The MIMIC-IV database

The MIMIC-IV database is an updated version of the MIMIC-III database [19]. It is a significant repository of anonymized health-related data, containing records of more than forty thousand patients admitted to critical care units at Beth Israel Deaconess Medical Center in America between 2001 and 2012. Within this database, there are approximately 1,400 patients related to hip arthroplasty. A comprehensive range of information is available, including demographic details, measurements of hourly vital signs taken at the patient's bedside, laboratory test results, medical procedure administered, prescribed medication, caregiver noter, imaging report and mortality data, covering both in-hospital and post-hospital discharge outcomes. The MIMIC-III and MIMIC-IV databases are crucial resources for conducting various analytical studies in fields such as epidemiology, improving clinical decision-making rules, and developing and evaluating electronic tools and interventions. This database is noteworthy for several reasons: It is freely accessible to researchers around the world, includes a large and diverse patient population from intensive care units, and provides detailed and frequent data [18].

Initial overview

As a part of the get-to-know-the-data set phase, I have created some graphs and visualizations to get an overview of the demographics of the MIMIC-IV database. In the following, gender, age, length of stay, BMI, and Charlson Comorbidity Index will be presented. For these general descriptions, only patients with hip arthroplasty surgery are included.

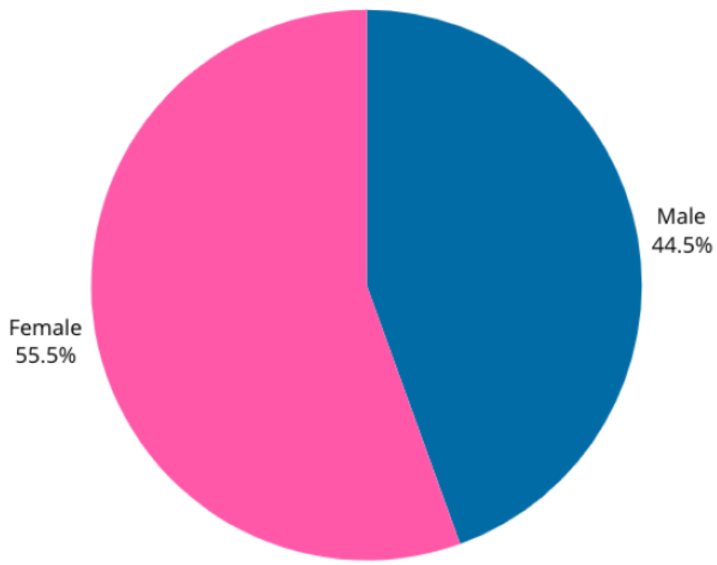

Figure 3.7: Gender distribution in the MIMIC-IV database

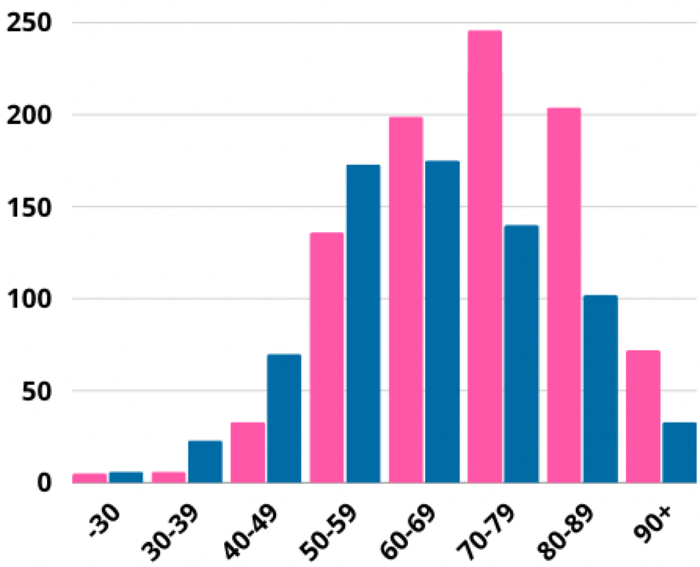

Figure 3.8: Gender and age distribution in the MIMIC-IV database

Figures 3.7 and 3.8 displays the distribution of gender and age within the MIMIC-IV database. Unlike the NAR database, there is a slightly higher representation of women in this database. In terms of age, the distribution is more scattered, but ones again more cases from 50 years and upwards. In addition, there is a higher number of cases for women in the age 70+ which can be good to note.

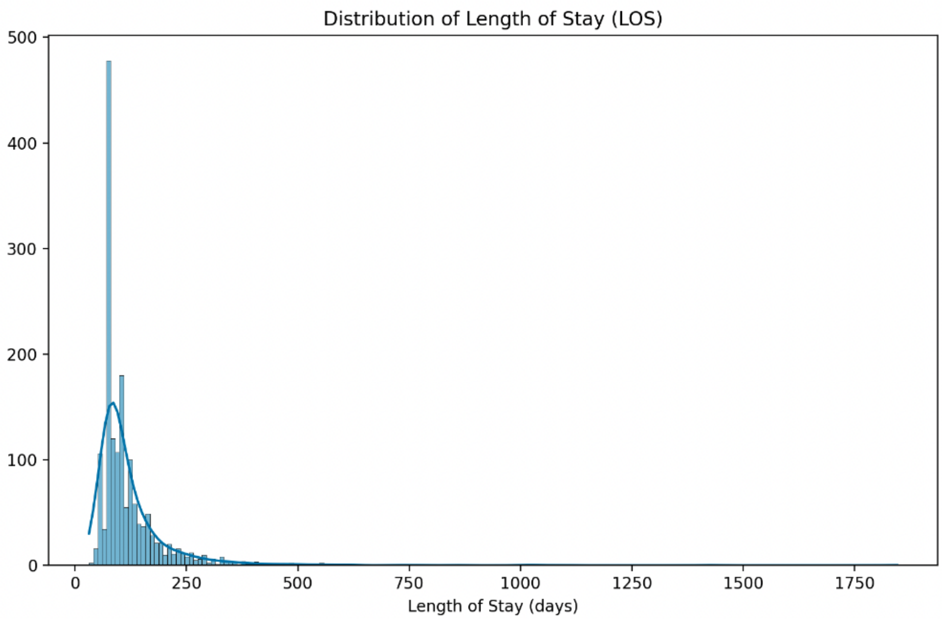

Figure 3.9: LOS distribution in the MIMIC-IV database [29]

Figure 3.9 shows a distribution of the length of stay a patient going through with hip arthroplasty has. It can vary a lot depending on recovery, or if the patient needs several surgeries. The distribution tells us that the patients in this database does stay for several days, mostly cases around 50-100 days, but there is also several cases from 250 days and more.

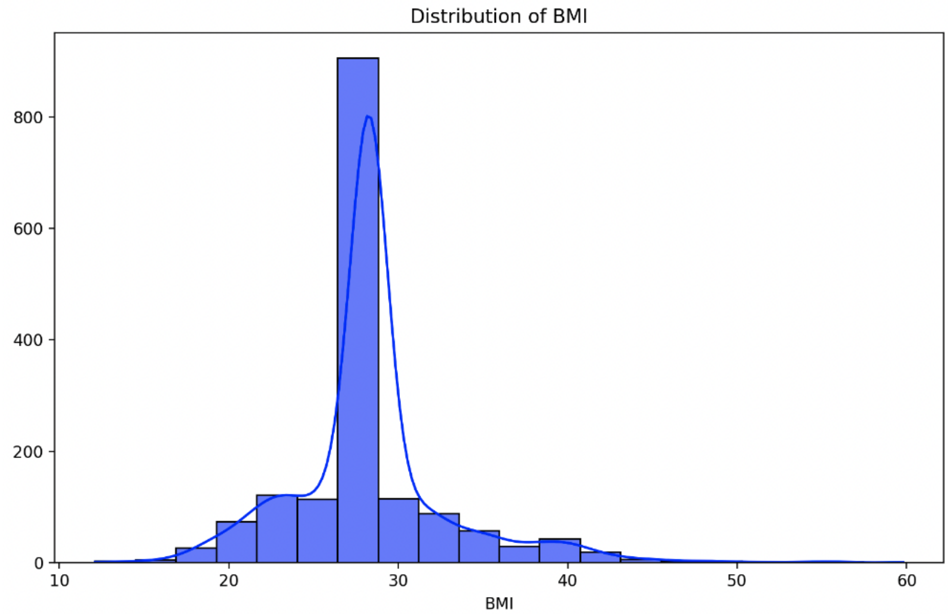

Figure 3.10: BMI distribution in the MIMIC-IV database [29]

The distribution in Figure 3.10 shows the Body Mass Index (BMI) of the patients. As seen most of the cases have a BMI of 28/29 which is categorized as "overweight". For context, a BMI between 18.5 and 25 is considered normal weight. A BMI between 25 and 30 falls into the overweight range, and a BMI over 30 is classifies as obese [33].

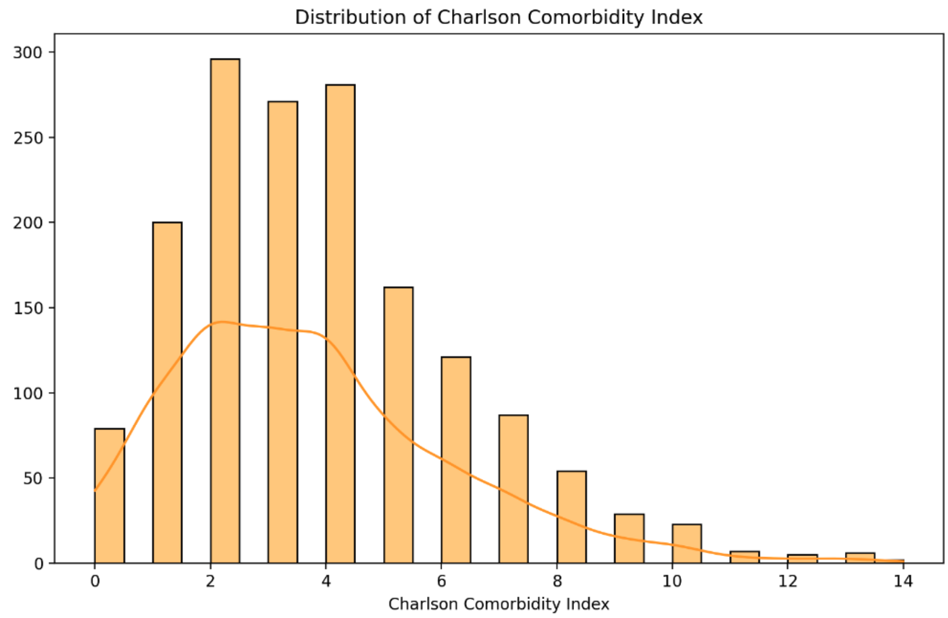

Figure 3.11: CCI distribution in the MIMIC-IV database [29]

The Charlson Comorbidity Index (CCI) is a widely used tool in healthcare to predict the risk of mortality and to assess the overall health status of patients based on their comorbid conditions. A score of zero indicates no comorbidities, while higher scores suggest a greater likelihood of mortality [21]. As seen in Figure 3.11 the majority of patients have scores below 5, indicating they have fewer comorbid conditions and therefore a lower risk of mortality and resource use. However, there are also many cases with scores ranging from 5 to 7, and some scores reach as high as 14. These higher scores represent patients with multiple severe comorbidities, indicating a significantly elevated risk of mortality and extensive resource needs.

Chapter 4

Methodology

This thesis aims to solve the challenge of digital twin visualization to help real users, such as physicians, researchers, navigationatients, understand the outcome of arthroplasty. Design science is a framework that solves such practical problems relevant for real environments and real users. It offers validated methods to create artifacts that are solutions for real-life problems. In this case, it is visualization that will bring better and more precise interpretation of digital twin and how it can help understand different clinical outcomes.
In this chapter, I am looking at design science, its principles, methods, and guidelines which I followed throughout this research. In addition, I present the concept of prototyping and conceptual design.

4.1 Design Science

Design Science is a problem-solving process that is grounded in theory, research, and systematic experimentation [27]. It focuses on developing and evaluating solutions to complex problems through a combination of creative thinking, scientific principles, and design iterations.
Hevner et al. [15] have developed both a framework and guidelines describing how to best carry out design science and research when creating new technology (Figure 4.1 and Table 4.1).
Figure 4.1 presents a conceptual framework for understanding, executing, and evaluating information system research combining the behavioral-science and design-science paradigms [15]. It illustrates the connection between rigor and relevance, two key factors in Design Science Research. Rigor ensures that the research is valid and reliable and contributes to the existing body of knowledge in the field. Relevance ensures that research provides valuable insights to organizations and professionals who can apply this knowledge to solve practical problems.

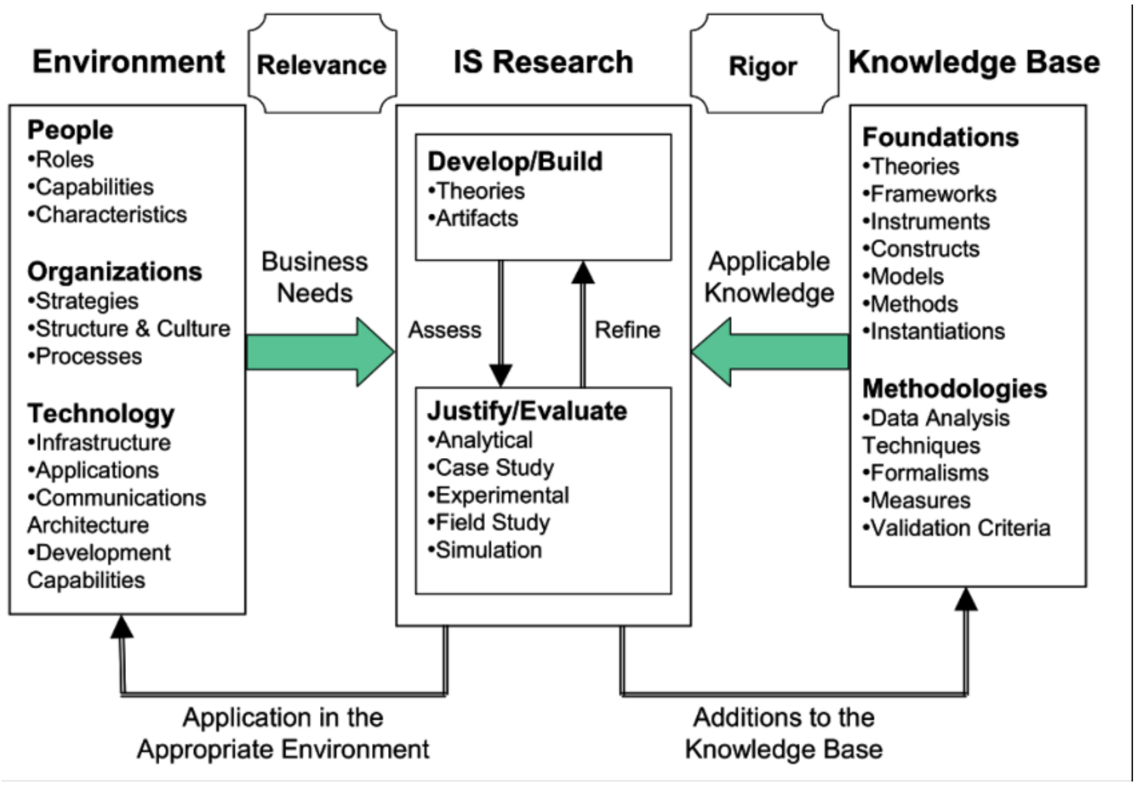

Figure 4.1: Information Systems Research Framework ([15], page 79)

Furthermore, Hevner et al. [15] propose seven design guidelines, summarized in Table 4.1. These guidelines are designed to assist researchers, reviewers, editors, and readers understand the requirements for effective and useful Design Science Research. To successfully create and evaluate an artifact, it is crucial to address each guideline in some way. However, there is no specific order in which they must be applied. The authors emphasize that researchers, reviewers, and editors should use their creativity and judgment to decide when, where, and how to apply each guideline to their specific research project.

Guideline	Description
Guideline 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.
Guideline 2: Problem Relevance	The objective of design-science research is to develop technology-based solutions to important and relevant business problems.
Guideline 3: Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
Guideline 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.
Guideline 5: Research Rigor	Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
Guideline 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.
Guideline 7: Communication of Research	Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences.

Table 4.1: Design-Science Research Guideline [15]

The research method that will be used for this thesis is Design Science. Design is used in this kind of research to refer to both the end artifact and the process that results in it [15]. It is mainly focused on problem solving and the innovations in technology, concepts, and goods that can enhance the analysis, design, management, implementation, and use of information systems. The existence of several paradigms does not negate the fact that the behavioral and natural laws that govern the domain still influence advances in design science. The reality is that new concepts and products depend on the researchers' prior understanding of behavior. The ability to implement, test, and adjust emerging ideas in a way that enhances and improves information systems and interactions with human behavior requires knowledge.

In "Interaction Design", Preece, Rogers and Sharp [27] explore the concept of design science for interaction design, which is a field concerned with creating meaningful and usable interactive products and services. They emphasize the importance of understanding user needs, behaviors, and preferences to inform the design process. By integrating theory, research, and practice, interaction designers can create innovative solutions that effectively satisfy user requirements.
One fundamental aspect of design science is the iterative nature of the design process. Iterations involve repeated cycles of design, evaluation, and refinement, enabling designers to progressively improve the quality and effectiveness of their solutions. Each

iteration helps designers gain insight into user needs and preferences, evaluate the usability of the design, and identify areas for improvement.
The authors also emphasize the role of prototyping in the design science process. Prototypes are tangible representations of design ideas that allow designers to test and gather feedback on their concepts. By creating prototypes, designers can simulate the interactions users would have with the final product, making it possible to evaluate factors such as usability, aesthetics, and functionality. Based on user feedback, designers can refine their prototypes, incorporating changes and enhancements to meet user expectations. This is important to remember, even though it is not always possible to implement everything at all stages of development or in all iterations. Prototyping is an iterative process and we can expect that in the higher (high-fidelity prototypes) iterations we should think about this type of user feedback.
Furthermore, "Interaction Design" [27] introduces various methods and techniques to gather user feedback, such as usability testing, interviews, and observations. These research methods help designers understand user needs, motivations, and mental models, facilitating the creation of products that align with user expectations. By engaging users throughout the design process, designers can validate their assumptions, uncover new insights, and make informed design decisions.
It is important to consider the context in which interactions occur. Designers must understand the environment, social aspects, and cultural factors that influence user behavior and preferences. By addressing these contextual aspects, designers can create inclusive and accessible solutions that meet a diverse range of users [27].

4.2 Prototyping

Creating a prototype and the process around it has been central for my thesis. When creating a prototype, there are several steps to consider [27]:

1. **Define the purpose:** Clearly identify the purpose and goals of the prototype. Determine what you want to learn or demonstrate through the prototype, for instance, to illustrate a design idea.

2. **Gather requirements:** Understand the requirements and specifications of the prototype. This involves considering user needs and other relevant factors that may influence the design.

3. **Design the prototype:** Create the initial design of the prototype. This can involve sketching or storyboarding. Determine the basic layout, structure, and functionality of the prototype.

4. **Choose the prototype fidelity:** Decide on the fidelity level of the prototype. Low-fidelity prototypes are quick and simple, focusing on the core interaction. High-fidelity prototypes are more realistic and often resemble the final product.

5. **Build the prototype:** Depending on the level of fidelity, build the prototypes using appropriate tools and techniques. Low-fidelity can be created using paper, sticky notes, or digital tools. High-fidelity may require coding or using specialized prototyping software.

6. **Test and iterate:** Conduct usability testing with the prototype. Observe users as they interact with the prototype and gather feedback on its usability, functionality, and overall user experience. Identify issues or areas for improvement. Based on the feedback, iterate on the design and make the necessary changes to refine the prototype.

7. **Repeat and refine:** Repeat the testing and iteration process, gradually improving the prototype with each iteration. The prototype should become more representative of the final product.

Prototyping is an essential aspect of the design process in "Interaction Design" [27]. The authors explain prototyping as a technique used to create low-fidelity representations of a design idea or concepts. These prototypes allow designers to gather feedback, test interaction, validate assumptions, and refine their designs before moving forward with the development process.
The authors emphasize the importance of prototyping as a means to communicate and explore design ideas. In addition, they highlight various approaches and techniques for prototyping, catering to different stages of the design process. They emphasize the value of low-fidelity prototyping, recommending methods such as sketching, storyboarding, and paper prototyping. These techniques enable designers to rapidly iterate and explore multiple design possibilities without spending excessive time or resources. On the other hand, a mid-fidelity prototype still has limited functionality, but in contrast to a low-fidelity prototype, this stage of the prototype includes interactive elements that demonstrate the navigation and interaction possibilities of the design [20]. Mid-fidelity prototypes are effective for validating interaction concepts by providing a clearer representation of the application's functionality compared to low-fidelity prototypes.
Furthermore, high-fidelity prototypes are detailed and interactive representations of the final product. It closely resembles the finished application in terms of design, functionality, and user interface [27].

4.3 Conceptual Design

Conceptual design is the initial phase of the design process where the broad outline and key characteristics of a project or product are defined [27]. In this stage, ideas are brainstormed, concepts are explored, and the overall vision of the project is established. This phase typically precedes detailed design work and focuses more on the big picture rather than specific details.
During conceptual design, multiple design alternatives are explored and evaluated. This involves considering different approaches, materials, or technologies that could be used to achieve the desired outcome.
In the early stage of the project, the data might be messy, incomplete, or not yet fully understood. Conceptualization allows designers to explore the data visually, even if the results have not yet been finalized. By creating rough visualizations, they can start to identify patterns, trends, and relationships in the data, which can inform further analysis and refinement. Visualizing data conceptually can help generate hypotheses or ideas about potential relationships or insights that the data might contain. Even if the data are not yet conclusive, creating visual representations can spark new insights or

lead to further investigation in specific areas. By creating a visual representation of the data, even if they are preliminary, collaboration among the team becomes easier in the context of conveying complex information and discussing potential interpretations or implications.

Conceptual visualization can help simplify and abstract the data, making it more understandable and manageable, even if the underlying complexity has not yet been fully addressed. It serves as a valuable tool for understanding and making sense of complex datasets, guiding further analysis, and facilitation collaboration between team members.

Overall, conceptual design sets the foundation for the entire design process by establishing the overall direction and vision for the project. Provides a framework for more detailed design work to follow and helps ensure that the final product or solution meets the needs of users.

Chapter 5

Methods

Methods in this project are focused on visualization aspects of both digital twin and clinical pathways. The work is concerned with visualization guidelines and takes into account various user groups, such as medical doctors, clinical researchers, and potentially even patients.
Two other master students, Sahlgaard [30] and Røise [29], in the project worked on the back-end part to provide knowledge about digital twin and clinical pathways using data mining in the NAR and MIMIC-IV databases. This means that the visualization was based on the results of the analysis and tools such as dashboards and infographics, which are suitable to bring forward the concept of digital twin. In this chapter, I present all these methods. The setup of the team work is also given.

5.1 Team-work

Within our team, we have consistently conducted regularly planned sprints and corresponding meetings to establish goals and objectives, as well as engage in discussions of the challenges and milestones achieved. To maintain a structured workflow, we have effectively used a Kanban board, allowing us to keep track and manage permanent tasks, ongoing assignments, and successfully completed achievements. This use of Kanban methodology has provided us with a clear and dynamic overview of our collective progress.
The project was divided into three different main paths, and our objective was to explore the intersection of technology, healthcare, and data visualization. The main paths were as follows:

1. **Digital twin creation:** Sahlgaard [30] has been responsible for creating digital twins for patients undergoing arthroplasty, by means of cluster analysis. His duties have included exploring the complexities of patient data, utilizing advanced modeling techniques, and converting medical nuances into a digital environment.

2. **Clinical Pathway exploration:** Røise [29] focused on mapping the clinical pathway of a patient undergoing arthroplasty. This involved a comprehensive exploration of the patient's journey from presurgical consultations to postoperative care. By understanding the intricate steps in the treatment of a patient, we have gained valuable insights that inform the development and refinement of our digital twins.

3. **Data Visualization:** The visualization task focused on translating complex data into visually comprehensible representations. The results of the cluster analysis have been given an appropriate conceptual visualization, and also clinical pathways were represented visually. A conceptual design was created to include both clusters and pathways.

5.2 Data mining

Data mining is a process of discovering patterns, relationships, and insights within large datasets. It involves various techniques and algorithms to extract valuable information from raw data [24]. In this project, both cluster analysis [30] and event logs [29] have been used, which have a direct impact on the visualization. The conceptual design now had two instantiatings within these two research paths. Although the final results have been included in the visualization to keep the concepts manageable, the visualization could include more details if the user evaluation suggested such solutions as beneficial.

5.2.1 Clustering

Clustering is an unsupervised learning method, which means that it does not require labeled data for training. Instead, it identifies natural groupings or clusters within the data based on the similarity of the data points. The goal of clustering is to partition the data into subsets, or clusters, where the data points within the same cluster are more similar to each other than the data points in other clusters [24].
Centroids play a crucial role in clustering algorithms, particularly in centroid-based clustering methods such as k-means. A centroid is the central point of a cluster, representing the mean or average position of all the data points in that cluster. In k-means clustering, the algorithm starts by randomly initializing a set of centroids. Then, it iteratively assigns each point to the nearest centroid and updates the centroids based on the mean of the data points assigned to each cluster. This process continues until the centroids converge, and there is minimal change in their positions between iterations. In general, centroids play a fundamental role in clustering algorithms, guiding the process of grouping similar data points together and extracting meaningful patterns from complex datasets.
The understanding is that similar clusters could be seen as the basis for defining twins. Based on the similarity of a new patient to various clusters, an assumption can be made about a possible outcome. If a new patient at hand seems to be similar to the cluster with a successful uneventful cluster of patients, then we could speculate that the patient twin is laying in the good outcome cluster. On the other hand, if the patient seems to be closer to the cluster with patients who experienced problems during or after surgery, it is likely that the digital twin is in the poor outcome cluster. This information is valuable when preparing for surgical treatment. The visualization should make clear connection between the patient and various digital twins by enabling one to look into the specifics of the clusters with good and poor outcomes.

5.2.2 Event Logs

Event logs are files that store information about significant actions or events in a computer system. The concept of event logs is universal across operating systems and devices. These files are chronologically ordered lists of recorded events, capturing important actions and occurrences [6].

5.3 Dashboard

A dashboard is an interactive control panel that includes various widgets such as sliders, checkboxes, and radio buttons. It features coordinated displays across multiple windows, showcasing various types of graphical representations such as bar and line graphs, heat maps, tree maps, infographics, word clods, scatterplots, and other visualizations [27].
In this thesis, dashboard is explored as a method to compile created visualizations, making them easily accessible and integrated into a single platform. Several dashboards were created to present different interactions with the content of digital twins, where content depended on results from various analyses, such as clusters and pathways. The advantage of dashboard is the possibility to combine these different methods in the same concept that allows two different aspects for any current patient at hand.
The design process with examples of the dashboard prototype will be presented in Chapter 7.

5.4 Wireframe

The initial conceptual models can be captured in wireframes, which are documents that outline the structure, content, and controls of a design [27]. These wireframes can vary in their level of detail and may represent either specific sections or the entire product. In this research, initial wireframes were made as a part of the iteration towards the final dashboard.

5.5 Infographics

An infographic is a type of visual information presentation that uses images, graphs, and charts with brief text to provide a clear summary. It is very helpful for explaining difficult concepts. The use of infographics is justified by their ability to quickly give a summary of a topic, explain complex procedures, display survey results or research findings, summarize long reports, enable comparison and support causes or problems [25].
In the context of the dashboard implementation, three key components are integrated:

1. **About the data / Patient demographics:** This section offers a visual breakdown of demographic information related to the patient population under analysis.

2. **Digital Twin:** Employing data visualization techniques, the concept of a digital twin is visually represented, capturing essential clusters and centroids to explain

patient characteristics.

3. **Clinical Pathways:** Utilizing infographic elements, clinical pathways are depicted, providing a visual representation of patient journeys through various stages of treatment or care processes.

5.6 Iterations

The idea of iterations within design science is given by Sharp et al. [27]. They emphasize that iterative design enables designers to incrementally refine their ideas and solution, providing opportunities for testing, feedback, and improvement throughout the design process. The iterative approach allows designers to learn from each cycle, incorporating user feedback, improving the design, and reducing the risk of errors. It also provides opportunities for designers to innovate, iterate, and converge toward a better design over time.
This research evolved through five iterations (Chapter 7), resulting in the mid-fidelity prototype that can be further developed through a new iteration into the high-fidelity prototype. The final goal of the project was to develop several conceptual designs. This incremental build-up of design to solve real environment challenges is an example of design science informed research.

5.7 Software and tools

5.7.1 Python

Python [28] has been used to understand and analyze the data. Two other team members have utilized Python in their research to a greater extent, while my usage was limited to using Python libraries such as scikit [31] and pandas [26]. I have used them to create graphs and demographics of the patients that have been later implemented within the conceptual prototype.

5.7.2 BigQuery

Google BigQuery is an AI-capable platform tailored for data analytics that empowers users to extract maximum value from their data. It is designed to support multiple engines, formats, and cloud environments [9]. This platform was used to work with and handle the MIMIC-IV database, in addition to the aforementioned Python libraries.

5.7.3 GitHub

GitHub is a web-based platform for developers to collaborate on projects and to easily work together on coding projects [13]. For this project, GitHub has been used to share code and results with each other. The repository is not included in this thesis due to patient security.

5.7.4 Figma

I have used Figma as the main tool for developing the prototype. It is a useful software for creating visualizations and prototypes. Figma has been used to develop a dashboard to gather the different visualizations in one place [12]. Figma also supports the development of complete prototypes by enabling the sense of navigation and interaction.

5.7.5 Canva

In addition, I have used Canva to create visualizations and graphs. Canva has a broad spectrum when it comes to creative visualizations and infographics [8]. Canva offers predefined templates and graphs that can be utilized visualization of the data, which again has been implemented in the Figma prototypes.

5.7.6 SPSS

Statistical package offering a great variety of multivariate statistical analyzes starting with summary statistics to decision trees, and cluster, discriminant, regression analysis to name a few mayor modules [17].

5.8 Evaluation

As stated earlier, two datasets were used, namely the NAR and MIMIC-IV databases, and results from the analysis on them have provided the basis for digital twins. It is up to potential users to make sense of the data and the results which is depending on the environment and content of the databases. The users can be evaluators in many ways, but of primary concern could be usability since the methods have their own measures of correctness and validity. The test of conceptual designs of the digital twin is a very new task and requires some reflection. Therefore, the cognitive walk-through has been used to process the concept of a digital twin and gain feedback.

5.8.1 Walk-Throughs

Walk-throughs provide a different route in the realm of design evaluation, offering a viable alternative to heuristic evaluation to anticipate potential user problems without the necessity of conducting formal user testing. As the term implies, walk-throughs entail navigating through a specific task or scenario with the product in question, actively observing and documenting any usability challenges encountered along the way. Unlike certain evaluation methods that directly engage users, walk-throughs rely primarily on the expertise of the evaluators and their understanding of usability principles to identify potential issues [27].
A cognitive walk-through is a method used in the field of design evaluation to assess the usability of a product or system from the perspective of potential users. Unlike some other evaluation techniques, cognitive walk-throughs focus specifically on understanding how users might think and interact with the product to accomplish their goals. The Cognitive Walk-Through will focus on two main aspects:

- **General Design Evaluation:** This involves an in-depth review of the visual and interactive elements of the digital twin visualizations. Key aspects such as layout, color schemes, navigation, and data presentation will be explored to ensure that they adhere to the principles of effective design and user experience.
- **Requirements Fulfillment:** Each visualization and (interactive) components will be assessed against the predefined requirements established. This includes evaluating whether visualizations provide accurate, relevant, and easily interpretable information to support clinical decision-making processes.

Chapter 6

Requirements

The ethical reflections considered throughout this research, the prototype requirements, and the initial personas are presented in this chapter.

6.1 Ethical concerns

Throughout the research process, appropriate steps have been taken to ensure that ethical concerns are taken into account and handled appropriately. Patient data in the research has been anonymized and used in a considerate way.
Sikt has approved this research. The approval can be found in Appendix A. To access the MIMIC-IV database, researchers must successfully complete the CITI Data or Specimens Only Research training, as an MIT affiliate; see Appendix B. The training includes completion of a course on safeguarding human research participants, incorporating the requirements of the Health Insurance Portability and Accountability Act (HIPAA). In addition, sign a data use agreement that outlines acceptable data usage and security standards. This agreement explicitly prohibits any attempt to identify individual patients.

6.2 Requirements

In the design process, requirements play a crucial role, as they serve as the foundation and guiding principles for developing a product or system. Requirements encompass both functional and non-functional aspects, each contributing distinctively to the overall design and functionality of the end design [27].
When designing a dashboard for digital twins of patients and medical data, there are several important requirements to consider. These requirements focus on providing a user-friendly and efficient interface while ensuring the privacy and security of patient data.

6.2.1 Functional Requirements

Functional requirements outline the specific functionalities and features that the product or system must possess to meet the intended purpose. They describe what the system is

expected to do in terms of input, processing, and output. These requirements act as the building blocks of the system's design. They guide developers in creating the necessary components, modules, and features to fulfill the user's needs. These requirements are typically tangible and measurable, providing a clear roadmap for development [27].
For this thesis, the selected functional requirements are:

- Present information about patient status
- Present information about the arthroplasty and its outcome
- Data analysis as a basis of digital twin definition
- Produce understandable data visualizations
- Include statistics whenever possible to provide healthcare personnel with details
- Differentiate visualization of patient-clusters and pathways
- Integrate different elements of visualizations in a dashboard

6.2.2 Non-Functional Requirements

Non-functional requirements are qualities or characteristics that describe how the system performs its functions rather than what functions it performs. They address aspects such as performance, reliability, usability, and security. They contribute to the overall quality and user experience of the product. Furthermore, they influence design decisions related to system architecture, performance optimization, and user interface design. Non-functional requirements are often critical to ensure that the system meets the desired standards and user expectations [27].
For this thesis, the selected non-functional requirements are:

- Aesthetics
- Flexibility
- Interoperability
- Testability

6.3 Personas

Sharp, Rogers and Preece [27] emphasize the utility of personas as a valuable tool for developers. Personas help developers understand diverse user characteristics and communicate objectives to both designers and developers. The primary goals of personas are to assist in making informed design decisions and to remind the team that real people will use the product. Rather than focusing solely on distinct user requirements, personas aim to provide a comprehensive understanding of potential user groups.
The authors define personas as "design fictions" that represent the needs, motivations, and behaviors of target users. These personas are based on qualitative and quantitative data from user research, such as interviews, surveys, and observations. By embodying

real user characteristics, personas enable designers to empathize with users and make informed design decisions.
The importance of personas in the design process is highlighted, emphasizing that understanding user needs and goals is fundamental for designing successful interactive systems. Personas shift the focus from designing for imaginary users to designing for real individuals with specific requirements.
Personas are valuable tools throughout the design process, helping designers prioritize features, make design decisions, and evaluate design solutions from different user perspectives. Using personas during ideation and prototyping ensures that designs align with user needs and expectations.
The authors also stress that personas should evolve over time. As user research continues and new insights are gained, personas should be refined to better reflect the user base. Keeping personas up-to-date is crucial as the design process progresses.

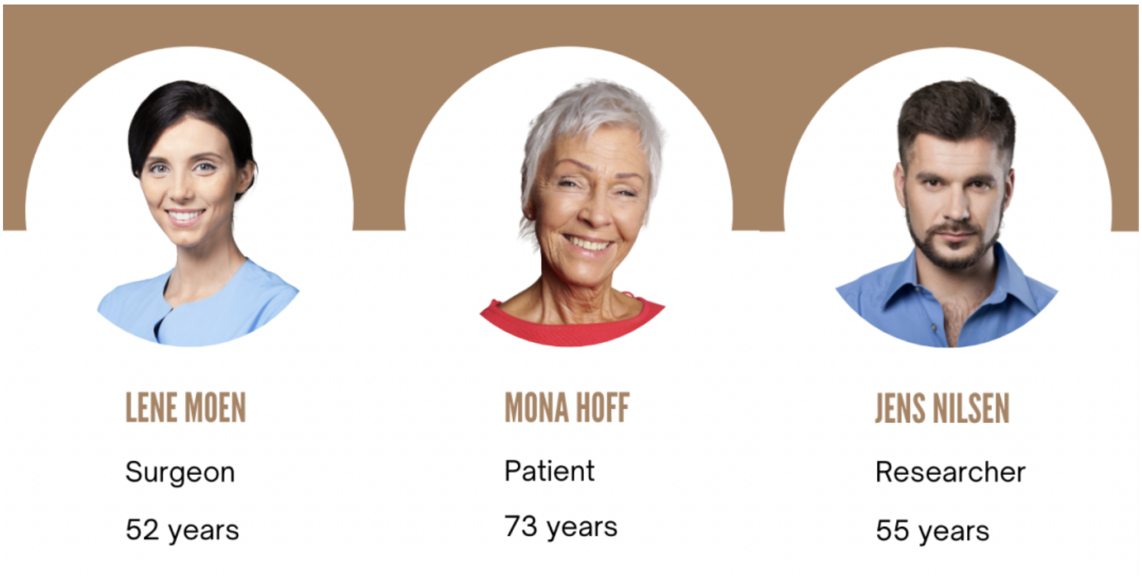

Figure 6.1: Initial personas

For this thesis, three initial personas were developed, each with a distinct background and a unique connection to the project's objective: one surgeon, one patient, and one researcher. These personas served as essential tools throughout the design process, enabling a deeper empathy with users and facilitating more informed design decisions. They were considered when prioritizing features, shaping the visualizations, and evaluating the solution.

Chapter 7

Visualization Development

7.1 Development Iterations

The final visualizations have been developed through 5 iterations. In this chapter, I am presenting details of each iteration, including the main feedback from walk-throughs done throughout the develobviousprocess.

Iteration	1	2	3	4	5
Dataset	NAR	NAR	NAR	MIMIC-IV	NAR MIMIC-IV
Fidelity	Low	Low	Low-Mid	Low-Mid	Mid
Method	Sketching, Figma, and Python	Figma	Figma, Canva	Figma, Canva, and Python	Figma, Canva, and Python
Evaluate	Cognitive Walk-Through with various experts				

Table 7.1: Overview of Iterations

7.2 Introductory Iteration

I am focusing on designing for all personas and addressing all their requirements. The initial visualizations are developed using clinical materials, aiming to create high-level visualizations that are understandable to all personas. I am aware that surgeons and researchers deal with detailed data while patients try to comprehend their own conditions. The goal is to conceptualize all important elements of visualization and introduce methods, always keeping in mind how the patient relates to these methods and to other patients. For instance, while a surgeon and researcher might focus on a patient's personal details and the expected outcomes, the patient is primarily concerned with understanding their current position based on their symptoms and what to expect. Therefore, I have mainly focused on creating several conceptual visualizations that could be of use for several personas, this was an attempt to find a good general conceptual design. The purpose was to figure out the concept of digital twin based on the patient data and results of the analyzes.

7.3 First Iteration

For the first iteration of the project, I have focused on exploring the NAR dataset and its associated demographic characteristics. The primary objective was to obtain an overview of the data, with particular attention to discerning any distinctive patterns or trends within the demographic variables. The results of these visualizations have been used in the initial overview of the database in Section 3.1.
To achieve this, visualizations were crafted utilizing ggplots and plotnine within the Python library. These visual representations not only facilitated a deeper understanding of the dataset but also serve as valuable tools for communicating insights. Additionally, Canva was used to improve the aesthetics of the presentation, ensuring clarity and impact.
I also started exploring the possibility of developing a dashboard to present the digital twins in the project. This dashboard will serve as a platform for showing information related to the digital twins and, thereby, enhancing the accessibility and interpretability of the insights from the dataset. In addition, it will make it possible to present both versions of the back-end data at the same time.
It is common to present data that are representative of a population or a group and create an understanding of the background and most interesting and significant variable that can quickly and directly convey essential data. The dashboard combines many times different types of data together and enables a more complex or deeper understanding. Usage of color is another important characteristic that enables understanding and often continuous understanding if colors are used consistently.

Low Fidelity Prototype

To test feasibility, I have created some low-fidelity prototypes using easy sketching by hand and Figma. Moving forward, these versions have been used to get feedback and develop more detailed design choices. The data used for this part are illustrative and will be replaced by real data in the following iterations.

Dashboard 1

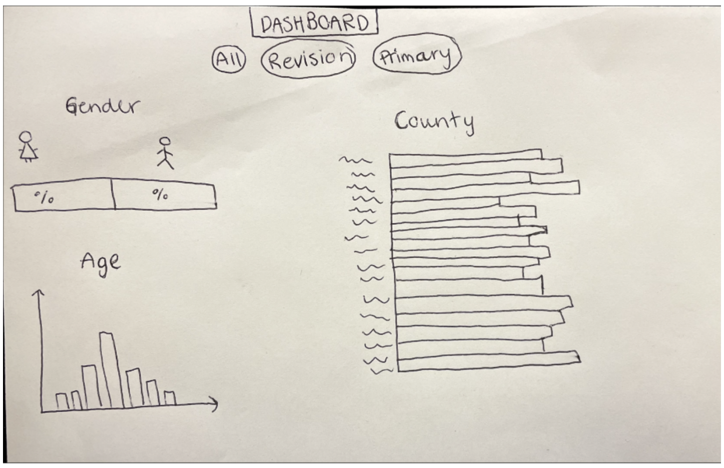

Figure 7.1: Low Fidelity Dashboard 1, sketch-version

Figure 7.1 differentiates between All, Primary, and Revision cases, highlighting the importance of Gender, Age, and County. Gender is shown as a percentage using symbols and bars for the sake of clarity. Age is represented in a bar graph, possibly divided by gender, with age groups on the x-axis and the number of people on the y-axis. The county data are displayed as a horizontal bar plot, with the county name vertically and the number of people horizontally.

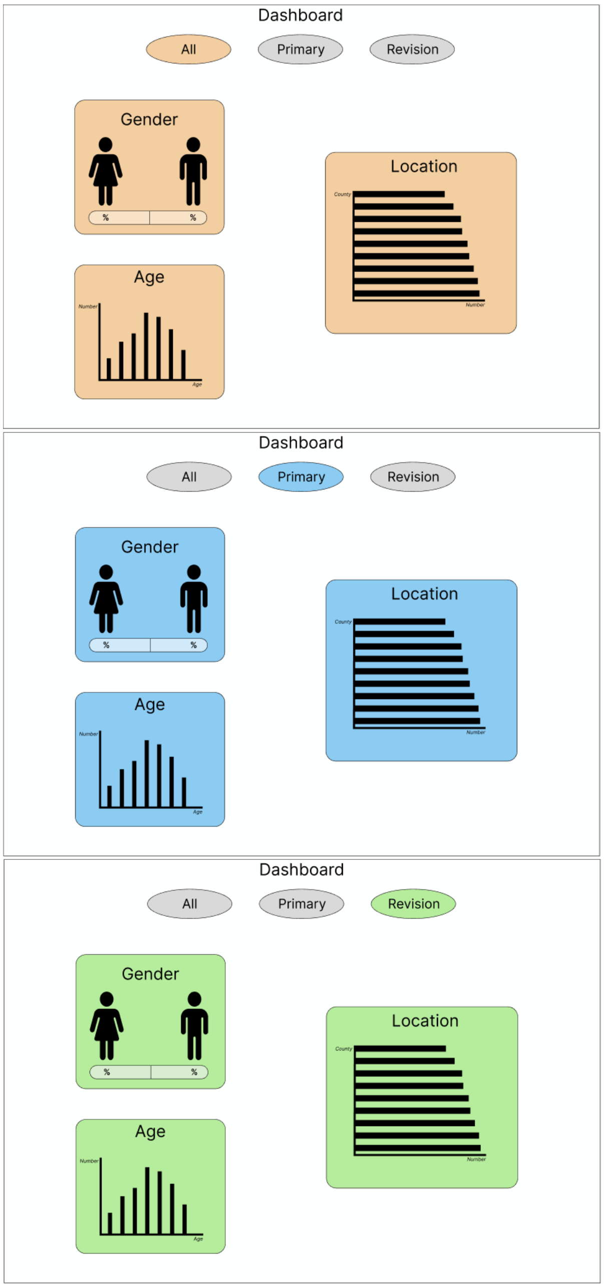

Figure 7.2: Low Fidelity Dashboard 1, Figma-version

Figure 7.2 consists of three stages, progressing from sketching to Figma. Simple visualizations have been created for the dashboard, utilizing colors for clarity and to separated the different section. The data is presented in a standard distribution format, with colors changing between the different data versions: All, Primary, and Revision.

Dashboard 2

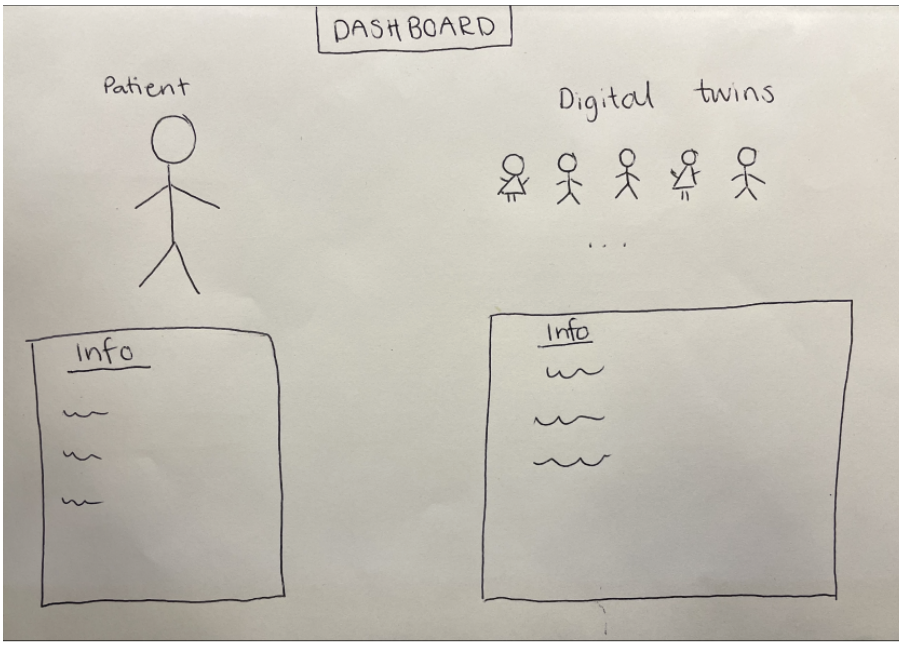

Figure 7.3: Low Fidelity Dashboard 2, sketch-version

Figure 7.3 represents one of the main ideas for the conceptual design. Includes patient information such as age, gender, county, and possibly name. Each digital twin of the patient is depicted as a small figure, with an information box underneath that represents the details of each digital twin.

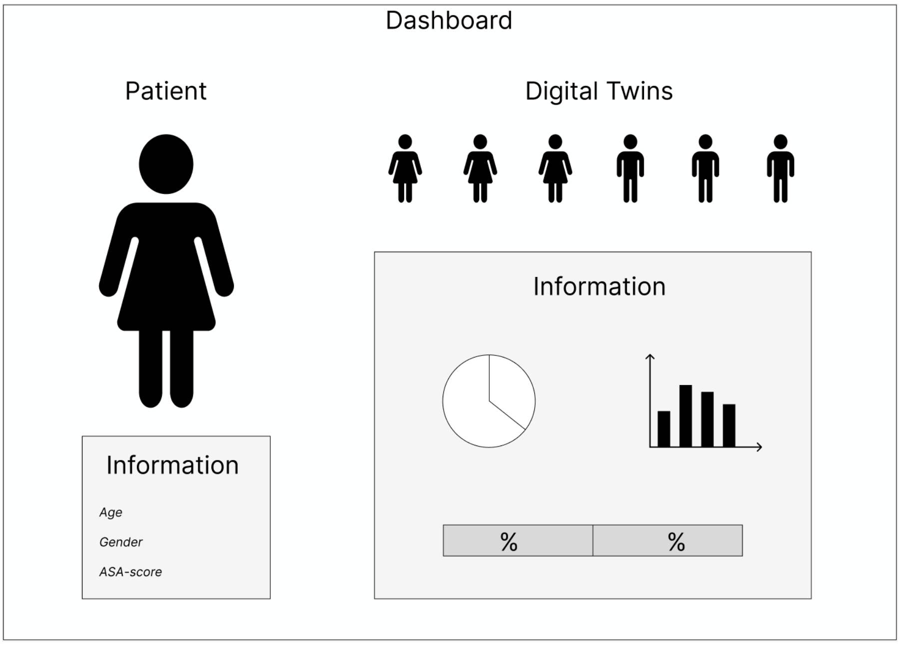

Figure 7.4: Low Fidelity Dashboard 2, Figma-version

Figure 7.4 is a Figma version featuring figures representing human beings, each accompanied by information boxes. The patient's box contains general information, while the digital twins' boxes include illustrations of various graphs resulting from the data analyses.

Walk-through feedback of First Iteration

The feedback received on the visualizations was overall positive. The suggestion was that, moving forward, it would be beneficial to specify and include information about the intended users, as different users may require different levels and types of information. The design should align with existing systems to facilitate easier implementation and increased likelihood of adoption by users. For now demographics and digital twins are divided into two sections, which is intended to be merged later on.

7.4 Second Iteration

In the second iteration, I have focused on further developing the content of Figure 7.3. Starting with an overview and including the intended content and divided into different sections as seen in Figure 7.5. Additionally, the data from the other two master students (Sahlgaard [30] and Røise [29]) in the project have been aligned with their results at this point.

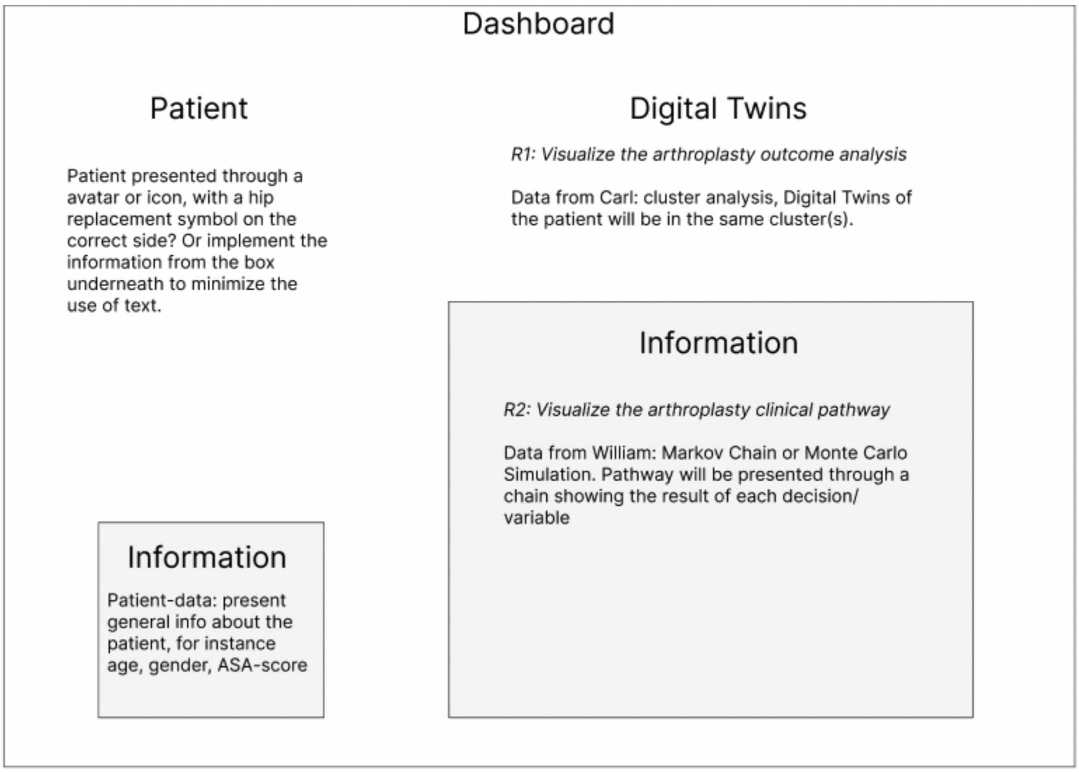

Figure 7.5: Overview of Low Fidelity Dashboard 3

More detailed design is following to include digital twin and clinical pathway (Figure 7.6).

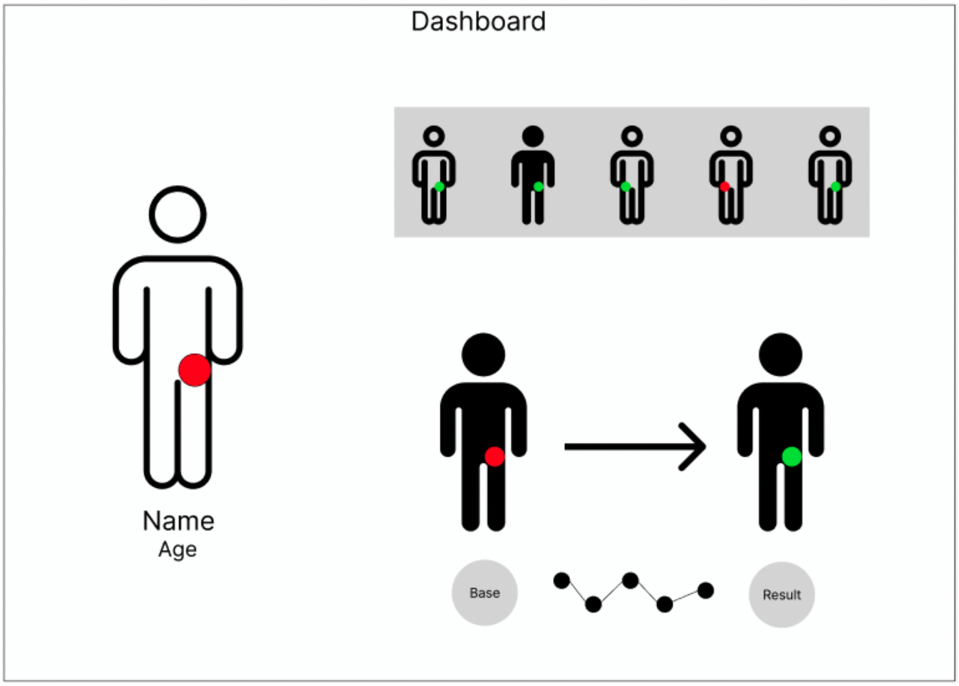

Figure 7.6: Low Fidelity Dashboard 3

Within the patient-section, patients are now represented through icons with an additional distinct hip replacement symbol on the correct side, this to enhance the user interface. Additionally, efforts were made for incorporating data currently displayed in a text box into more visually intuitive formats.
The presentation of digital twins aligns with the exploration of RQ1 - *How to best visualize the arthroplasty outcome analysis?* with the intention of using it with the research carried out within the project, and thus contributing to a more insightful presentation of the results of the arthroplasty.
The information section addresses RQ2 - *How to best visualize the arthroplasty clinical pathways?*. This section will be connected to the research on the pathways also carried out within this project.

Walk-through feedback on Second Iteration

Figure 7.6 has no subheaders, and the title should be more descriptive than just "Dashboard". Although the different sections are clearly divided, each section should include additional explanatory information.

7.5 Third Iteration

In the third iteration, emphasis is placed on visualizations of both cluster analysis and clinical pathways of the NAR dataset.

Cluster Analysis

Visualization of cluster analysis is expected to reflect the digital twins of the patients, or perhaps more precisely the digital cousins. Drawing on Sahlgaards' contributions

[30], patients are categorized into clusters, each characterized by a centroid denoting its central tendencies, using two or several variables. Although not yet achieving correct digital twin status, the result of data analysis have shown that clusters could not be fully distinguished from each other, so it was hard to conclude the presents of digital twins. The lack of separation suggested that another concept could be introduced, such as digital cousins. Cousins share many similarities; however, they are not well-defined digital twins. Regardless of this, the conceptual design of the twins will still indicate separate clusters while using the instantiating (real examples), which will make the clusters overlap.

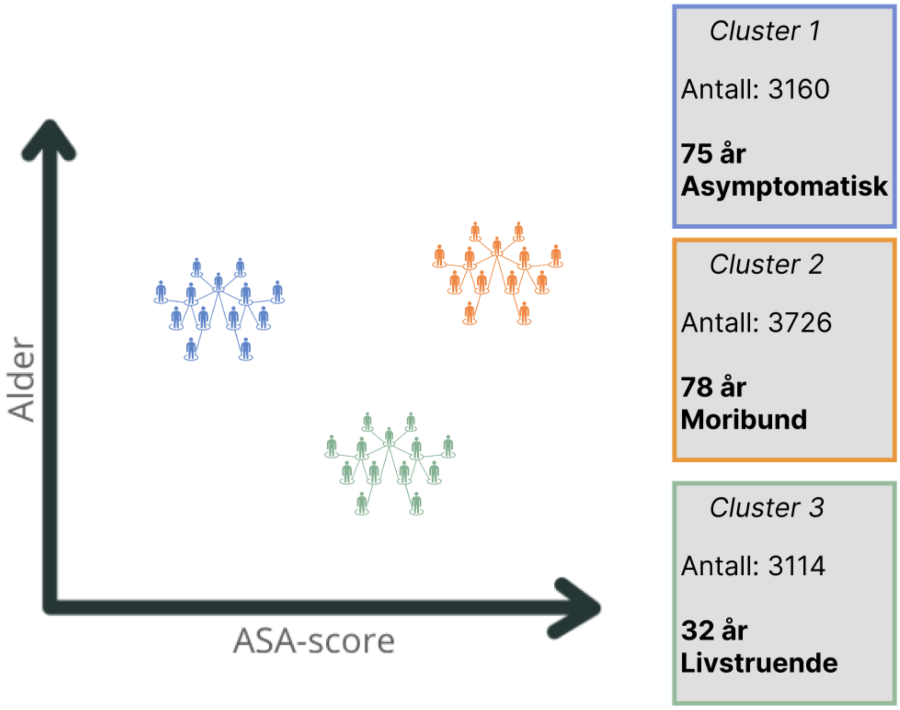

Figure 7.7: Visualization of Cluster Analysis NAR

The data displayed in Figure 7.6 is based on a initial cluster analysis of the NAR database, with random centroids for age and ASA-score, as the final cluster gave almost identical results for the three clusters. Complications was also used for the clusters. The system counts the average number of complications connected to the cluster, which is not informative, as the list is not ordered according to the risk or severity of the complication. If this was the case, then the average number would suggest how severe the complications are seen in a particular patient cluster. However, using ASA score can help mitigate this problem, since it is an ordinal number that goes from low to high risk.
As a contribution to this stage of the datamining, discriminant analysis was performed to determine which variables could be significant. The results of the analysis indicated that the patient groups could not be efficiently separated and the notion of a digital twin was not easy to derive.

Clinical Pathway

Furthermore, attention is directed towards visualizations of clinical pathways for patients. One key consideration is the optimal presentation of event logs within these pathways. Drawing from Røises' research [29], a generalized pathway is constructed based on the dataset, connecting the sequential steps patients are likely to traverse. These pathways encompass common stages shared by all patients, as well as divergent routes required by additional tests or interventions. Figure 7.8 shows the results provided by Røise [29], following the visualizations made in an attempt to create a user-friendly clinical pathway.

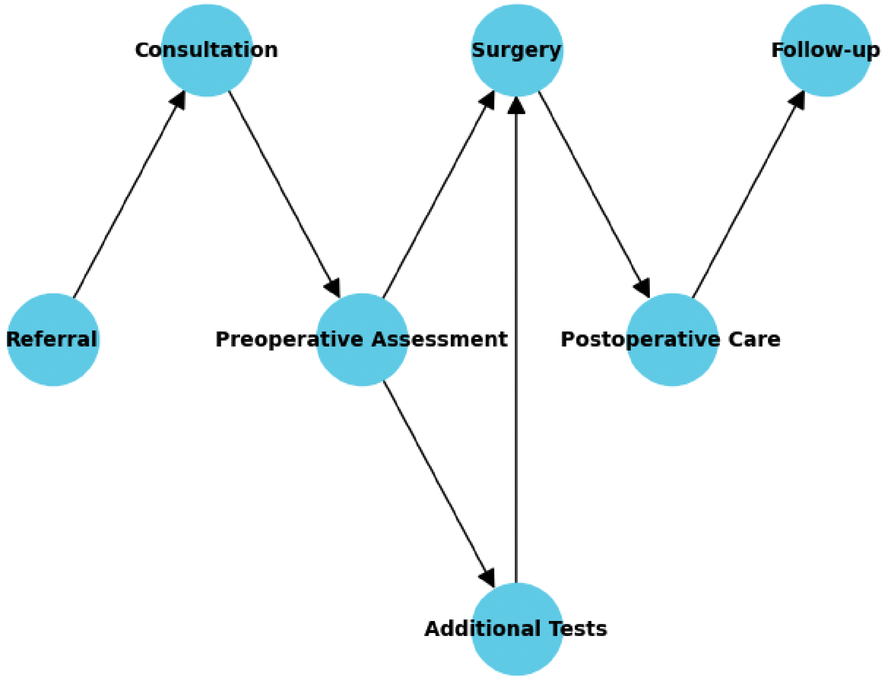

Figure 7.8: Clinical Pathway in the NAR database from Røise [29]

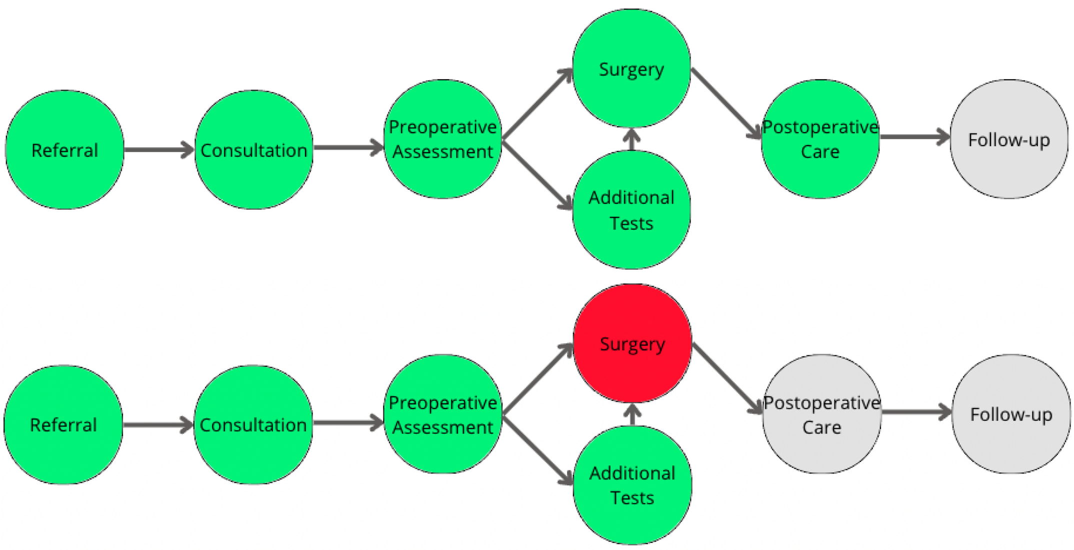

Figure 7.9: Example clinical pathways in the NAR database

The clinical pathways in Figure 7.9 are represented with circles named after which state the patient is in, color to identify positive or negative outcome, and arrows to show the order.

Walk-through feedback on Third Iteration

The cluster analysis visualization currently uses a mix of English and Norwegian; it should be entirely in English, with the option to switch to Norwegian implemented later. Furthermore, the arrows in Figure 7.7 are misleading as it is unclear whether the graph continues. For the clinical pathway, the use of colors is clear, but more information is needed. It is not clear what determines whether a point has a positive or negative outcome. Furthermore, additional information should be provided about the different steps of the pathway.

7.6 Fourth Iteration

In this iteration, the MIMIC-IV dataset is explored and used to develop a dashboard to present demographics, cluster analysis, and clinical pathways. The dataset includes richer clinical data, covering demographic information, procedures, and diagnoses related to various hospital stays for patients. The initial step involved getting acquainted with the dataset by examining its fundamental aspects and the demographics of the included patients. This helped in understanding the scope and structure of the available data.

Wireframe

For the general idea of a dashboard, a wireframe was created as a framework for future design.

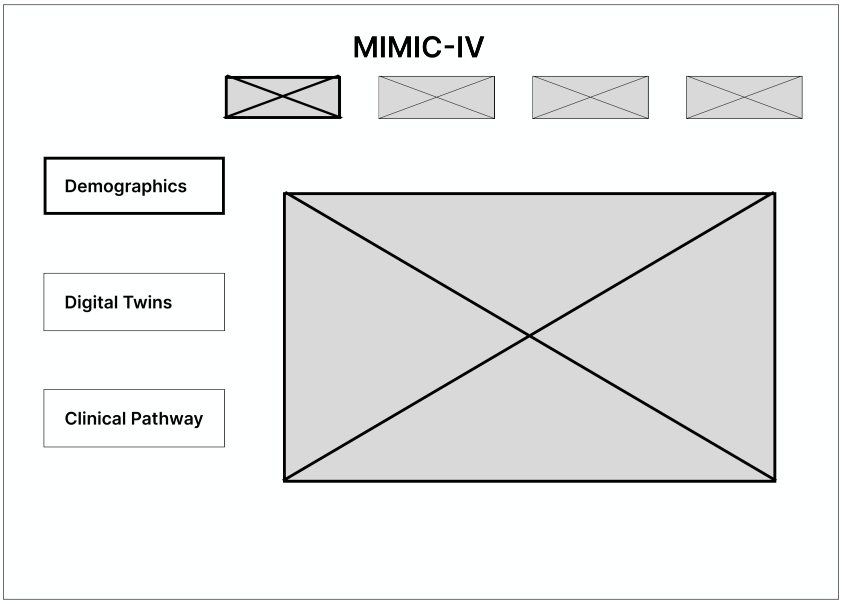

Figure 7.10: Wireframe of Dashboard of the MIMIC-IV database

Figure 7.10 is split into three main sections, "Demographics", "Digital Twins" and "Clinical Pathway". The user can easily maneuver around the page and choose which section or combination of sections they would like to see.

Cluster Analysis

In the MIMIC-IV database, the clinical picture is given in great detail, and there are many variables describing patient status, starting with diagnosis data, continuing to procedure data and including data about difficulties (comorbidities). Such a database implies that the clusters will be different from the ones calculated for the NAR database. It could be assumed that the number of clusters and significant variables used in the calculations will provide a different picture resulting in different definitions of digital twins. Therefore, the main focus has been on how to explore the MIMIC-IV database and reason about the possible number of clusters and their visualization.

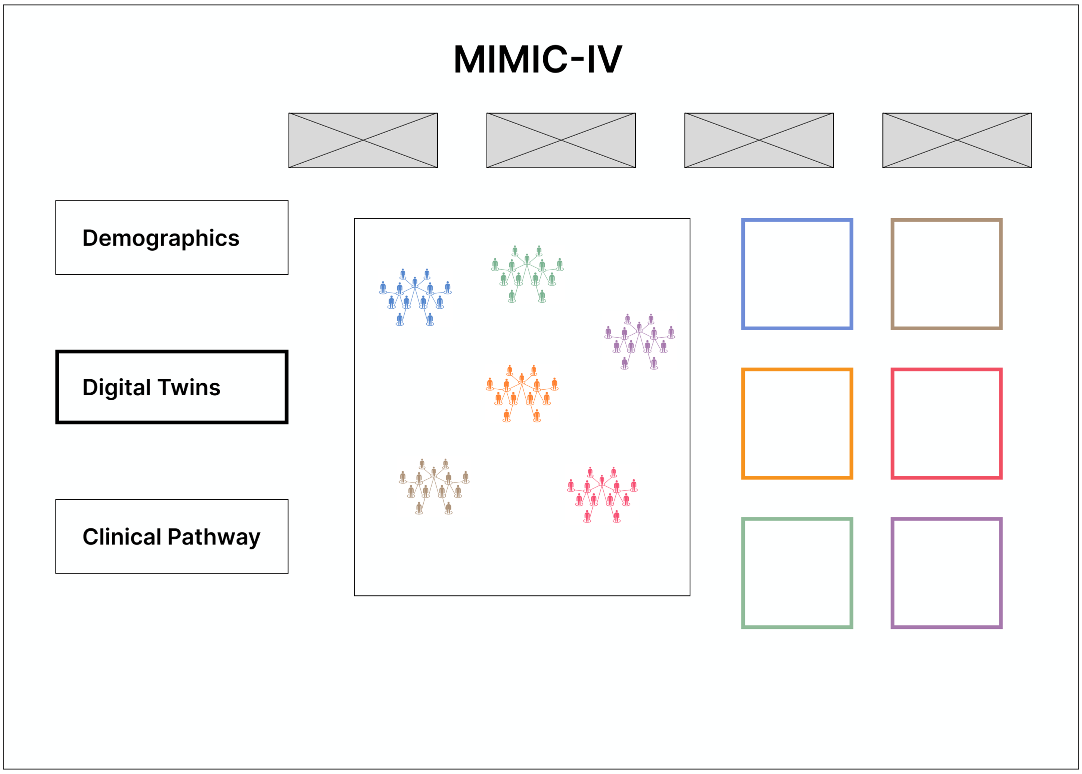

Figure 7.11: Cluster analysis example for MIMIC-IV

The section for Digital Twins shows six clusters distinguished by different colors. As the number of clusters can vary, the framework shows an example of six clusters and relating information shown in boxes with the same color. Wireframe boxes are placed on the top to suggest the possibility that users can decide which variables to include in the cluster analysis.

Clinical Pathway

Subsequently, efforts were made to construct the clinical pathways by leveraging the detailed clinical data. This involved linking diagnoses and procedures to specific times during hospital stays for patients. Although diagnoses were generally established at the time of admission, procedures could be placed at specific points after admission (Figure 7.12).

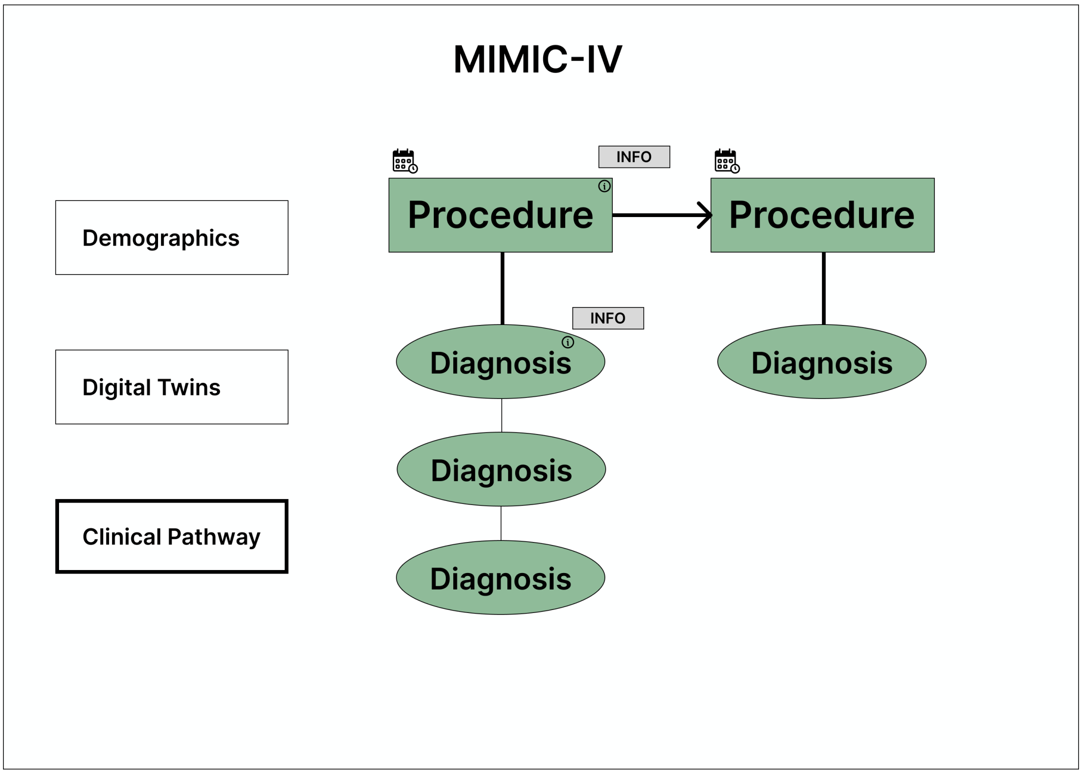

Figure 7.12: Clinical Pathway example for MIMIC-IV

Square boxes identifies procedures while diagnoses are placed in circles. As each box only can contain a short description or title, an information box is added to give the opportunity to complement with an expanded explanation. In addition, a calendar is placed on top of the procedures to point to where the timestamp would appear.

Walk-Through feedback of Fourth Iteration

Observations found during the walk-through emphasized the need to make the heading and the "main" menu of the dashboard more prominent. Additionally, clearly differentiate between the separate parts of the dashboard so that it is obvious which actions or buttons are connected to which information.

7.7 Fifth Iteration

This section presents the final design for the presentation of demographics, clusters, and clinical pathways for both the NAR and MIMIC-IV databases.

7.7.1 The NAR database prototype

In spite of the differences between the NAR and MIMIC-IV databases, I am suggesting similar concepts for both databases. However, I need to acknowledge the specifics of the NAR database in terms of variables, scope of database, and number of patients. It could be assumed that the great number of patients could provide a very interesting analysis, but at the same time, I had to acknowledge that this database is simplified compared to the full electronic patient record. Regardless of that, cluster analysis is a useful tool to apply when we want to learn and generate hypotheses about patients and their surgical results.

Demographics

Demographics in NAR describes the basics of patient data. We can see that age, gender, and county are the main general descriptor that is relevant for a national registry, such as the NAR database. I am showing some variations of the dashboard that can be offered to the users.
In Figures 7.13 and 7.14 the main header and menu are positioned on the left side, with an additional menu running along the top. The selected buttons on the menus are highlighted with bold text and thicker lines around them. The headings of the subsections have been added for clarity. The figures are similar, but details matter to users, so they are still offered two designs to choose from.

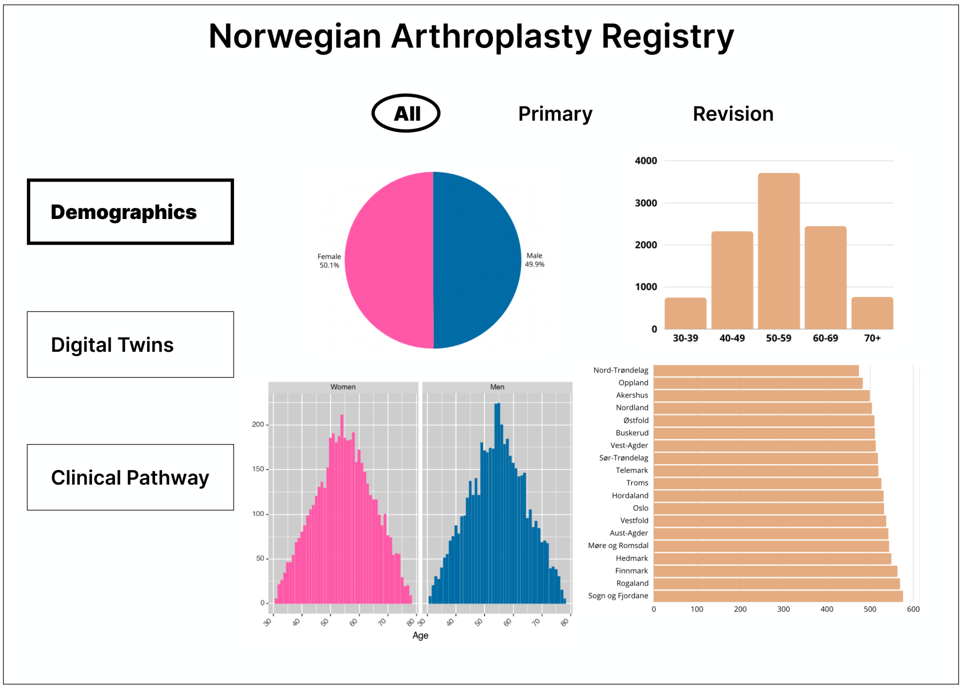

Figure 7.13: "Demographics" dashboard NAR version 1

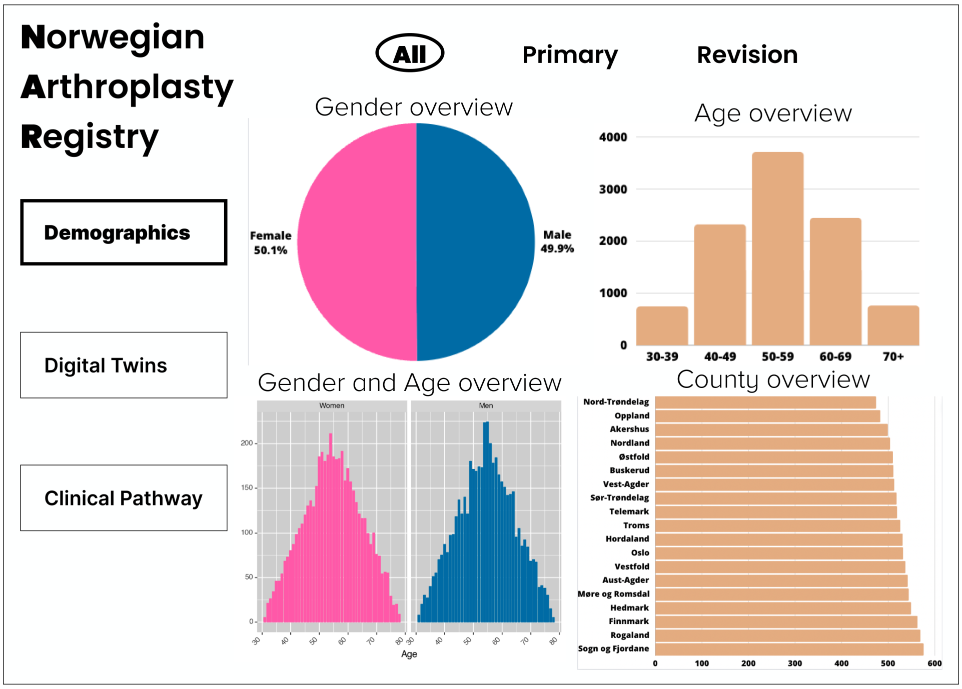

Figure 7.14: "Demographics" dashboard NAR version 2

Cluster Analysis

Here I am presenting the initial cluster, final cluster, and an example of the cluster resulting from a data analysis. This makes the whole development transparent and shows how to use the concept with one example of instantiation.
Figure 7.15 displays the initial clusters, which are used to clearly distinguish between the groupings.

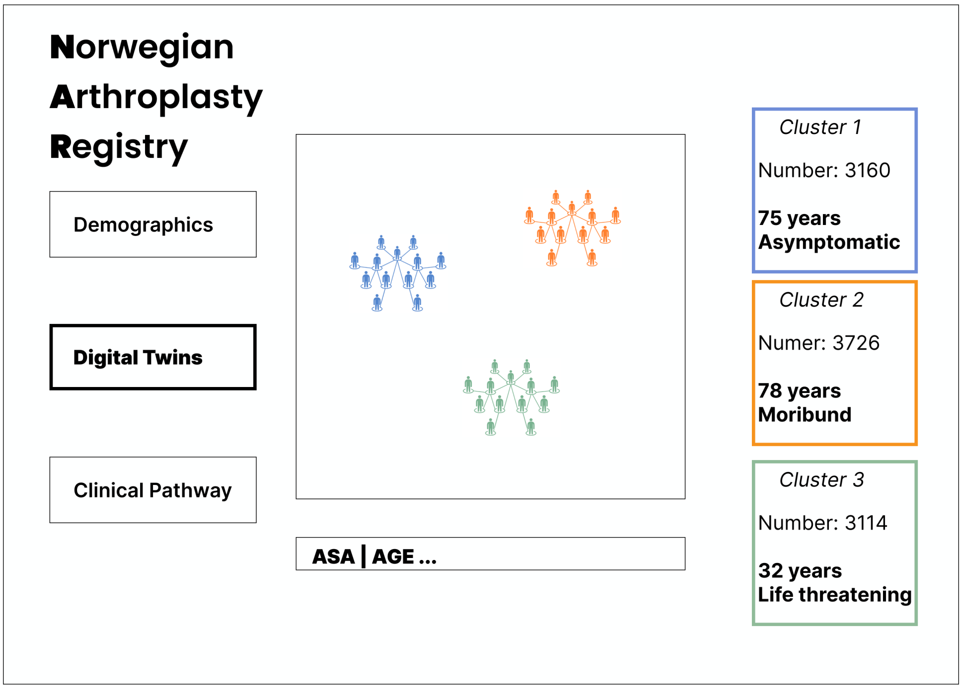

Figure 7.15: "Digital Twins" dashboard with initial clusters NAR

Furthermore, an example is provided to illustrate how the final cluster results for the NAR database look.

Final Cluster Centers

	Cluster		
	1	2	3
Pasientens alder i år ved første operasjon	54,70	54,82	54,33
Peroperativ komplikasjon, spesifisert (primær)	16	92	51
ASA klasse (primær)	4	4	4

Figure 7.16: Example of final clusters in the NAR database

Figure 7.16 presents clusters as a table with three centroids which is a form often seen in statistical analysis, as done in Sahlgaards' thesis [30].
Figure 7.17 is a general design that includes data from real calculations (Figure 7.16).

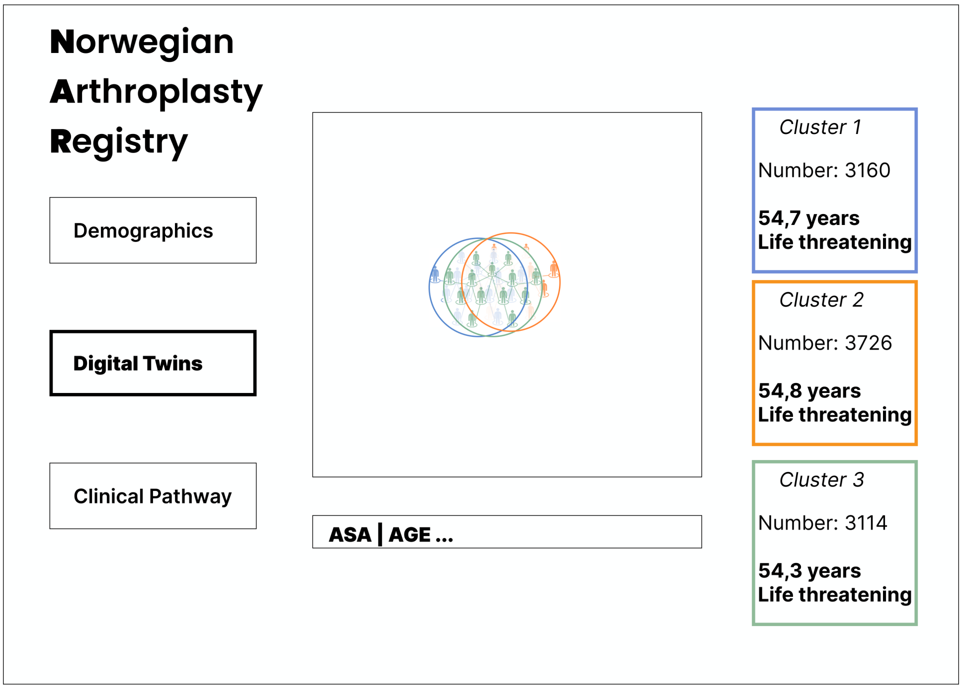

Figure 7.17: "Digital Twins" dashboard with final clusters in the NAR database

Figure 7.17 shows the final clusters, which in this case are almost identical. In addition, the information boxes containing the centroids are presented on the right side, along with a field for the variables on which the clusters are based.

Clinical Pathway

The visualization in this subsection demonstrates the application of clinical pathways within the context of cluster analysis.

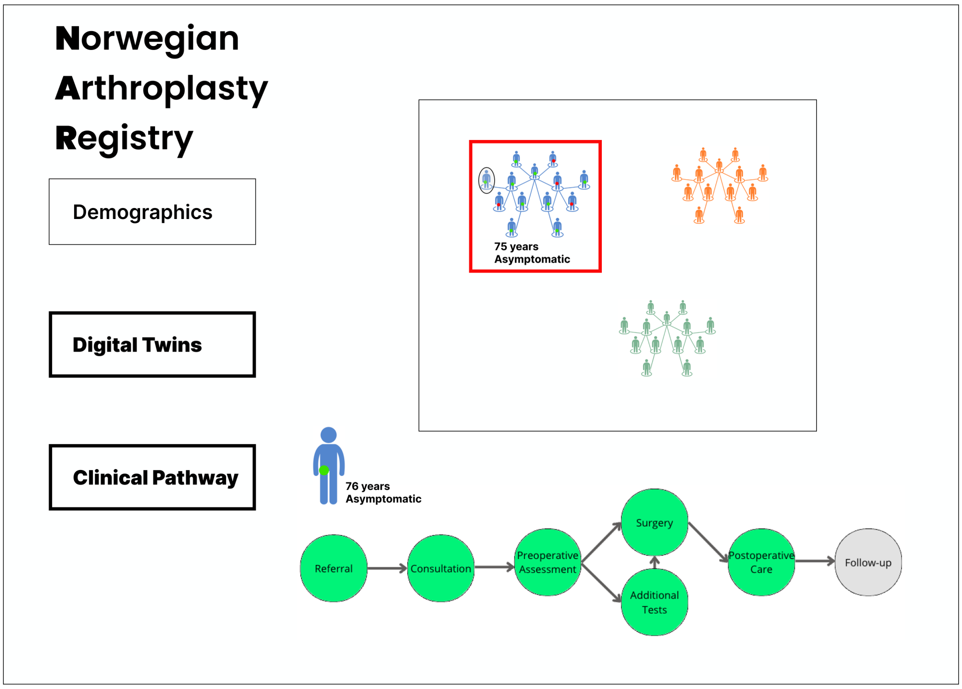

Figure 7.18: "Clinical Pathway" and "Digital Twins" dashboard NAR

Figure 7.18 presents a detailed view of the integration between clinical pathways and cluster analysis visualizations. The selected cluster is highlighted with a red box, with each patient marked by a green or red indicator to represent their operation outcome. In the following, the clinical pathway of a patient chosen from the cluster is displayed with its own set of visualizations.

7.7.2 The MIMIC-IV database prototype

The MIMIC-IV database differentiates from the NAR database. There is a lot of information on clinical pathways in addition to demographic data. The most specific feature is that all diagnosis and comorbidities are described using icd codes. In some cases, this means a huge number of codes attached to one patient and different hospital stays. This makes analysis and visualization more demanding than in the case of the NAR database.

Demographics

The demographic data are comprised of more clinical findings and there is no county data as seen in the NAR database. New interesting variables are Charlson Comorbidity index and BMI, which are specific to the MIMIC-IV database and therefore the graphical visualization is different from the NAR database as seen in figure 7.19.

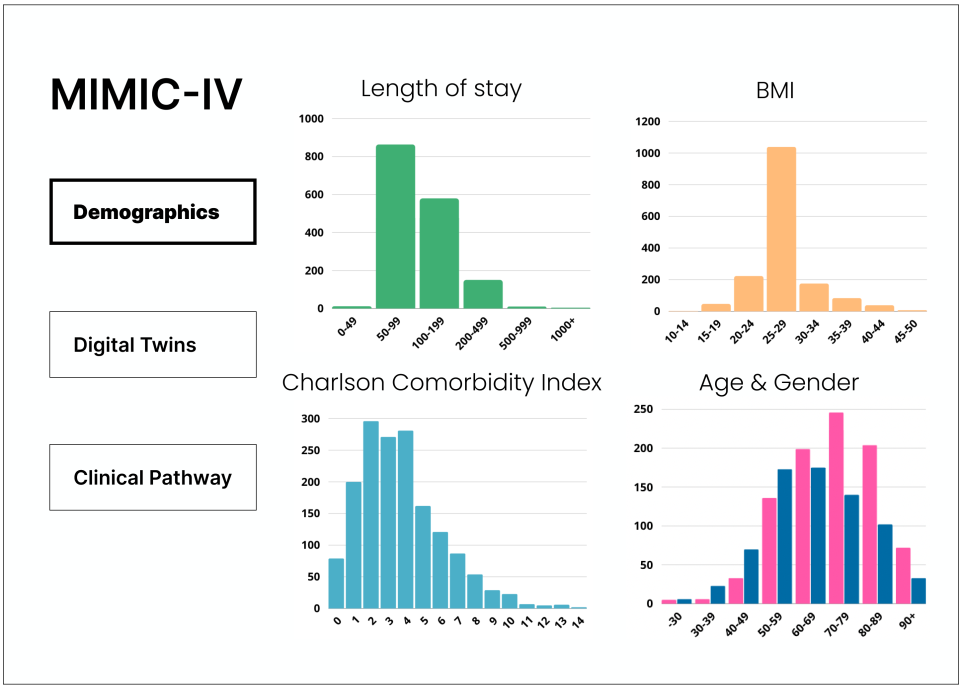

Figure 7.19: "Demographic" dashboard MIMIC-IV

Cluster Analysis

This subsection presents the framework for cluster analysis alongside an illustrative example derived from a Python-based cluster analysis.

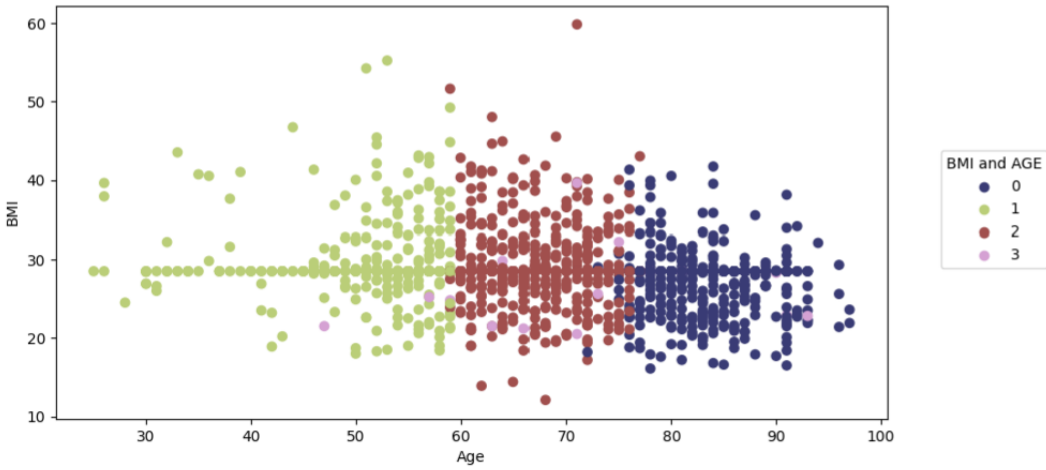

Figure 7.20: Example Clusters in the MIMIC-IV database from Sahlgaard [30]

Figure 7.20 is an example on how clusters can look like from analysis in Python. Three bigger and clearly separated clusters. There is also a minor fourth cluster represented with yellow dots that is included in the three larger clusters.
The following are two examples on how these kinds of results could be visualized in a dashboard.

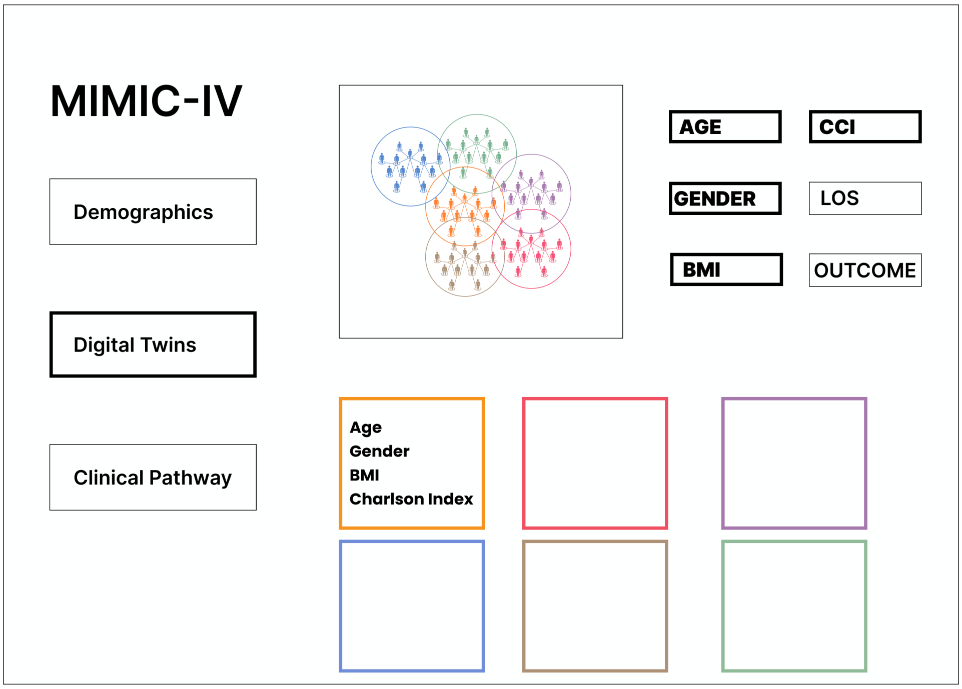

Figure 7.21: "Digital Twins" dashboard MIMIC-IV version 1

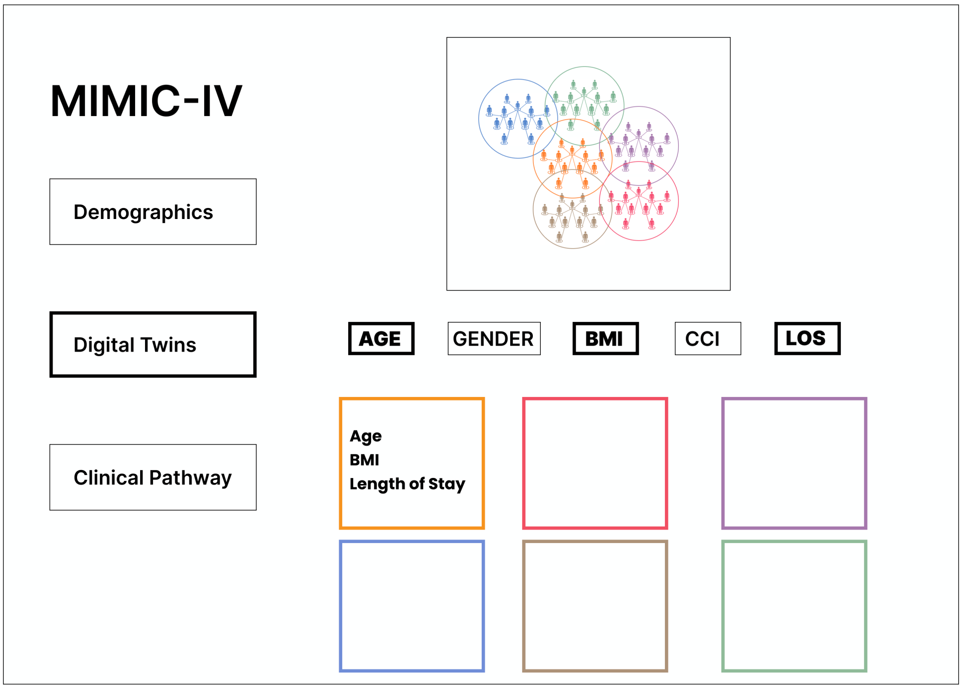

Figure 7.22: "Digital Twins" dashboard MIMIC-IV version 2

Figures 7.21 and 7.22 include menu buttons to select variables to be included in the clustering analysis displayed. Additional information boxes are provided to present the centroids for each cluster.

Clinical Pathway

For the clinical pathway in the MIMIC-IV database, the data are richer and therefore implemented only to show how this could be visualized. Using results from Røise [29], an attempt is presented to create a framework with some examples from the actual database.
In contradiction to the NAR database, the clinical pathways for MIMIC-IV are more personalized for each patient. A hospital stay is related to several procedures and diagnosis. Procedures can be defined to a specific time, while diagnosis is now placed at the time of admission. Figure 7.23 is an example of a patient stay, where circles indicate diagnoses, while squares are procedures. In addition, arrows are placed between, but they only indicate the timeline for procedures, while the diagnoses are all set at the same time. This is an example of a patient for whom the clinical pathway is relatively simple. Other more complex pathway examples can be found in Røises' thesis [29].

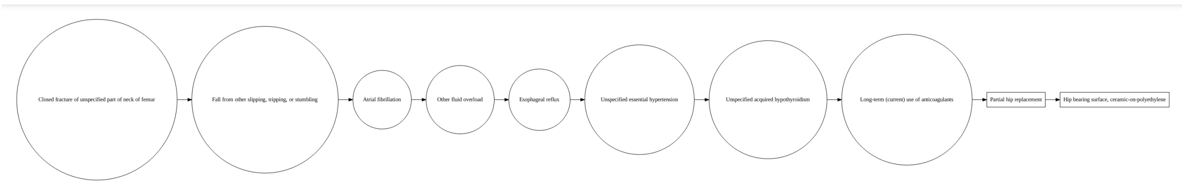

Figure 7.23: Clinical Pathway example MIMIC-IV from Røise [29]

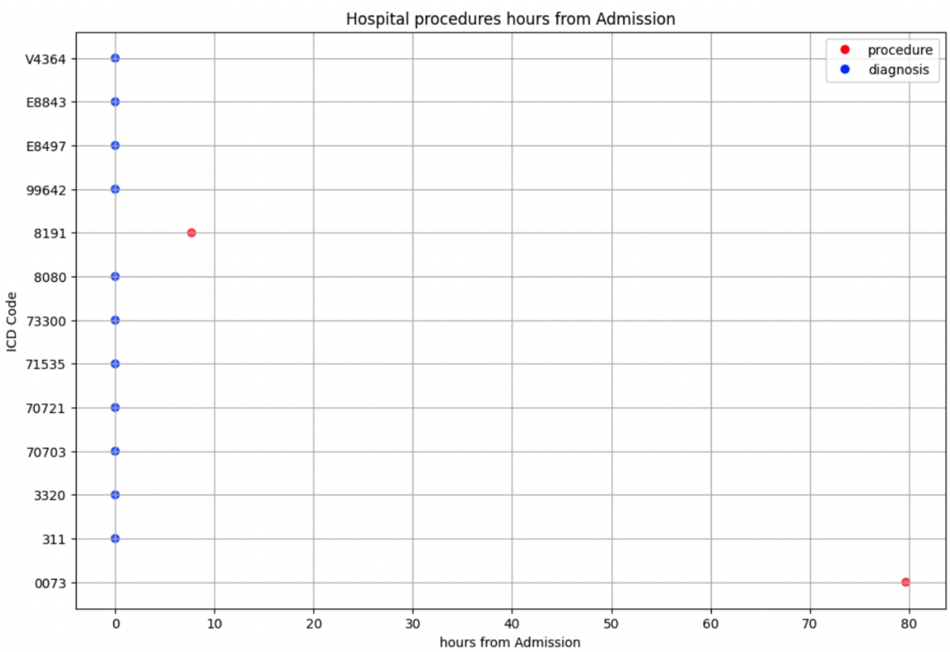

Figure 7.24: Procedure and Diagnosis example MIMIC-IV from Røise [29]

The figure 7.24 shows a more detailed picture of procedures and diagnoses given an event log timeline. This is an instantiation of the event shown in the clinical pathway dashboard, as shown in Figures 7.25 and 7.26.

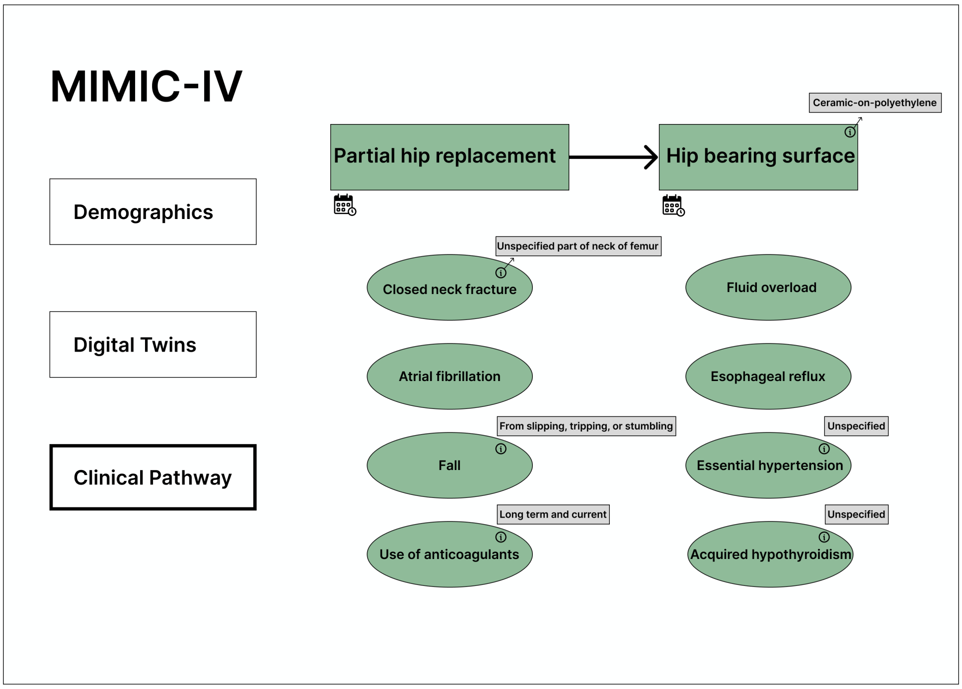

Figure 7.25: "Clinical Pathway" dashboard MIMIC-IV version 1

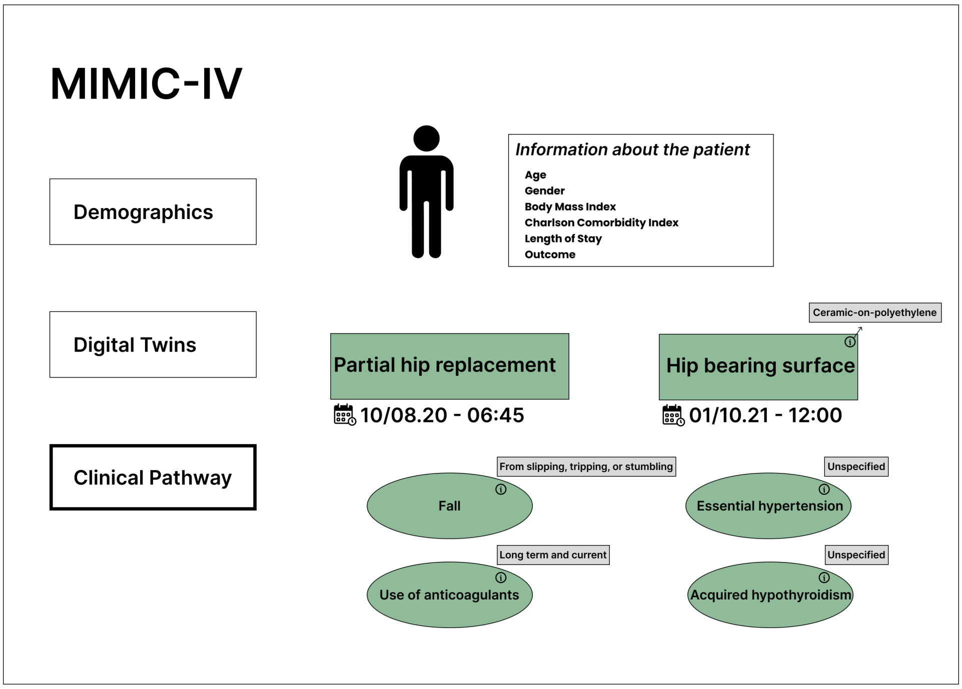

Figure 7.26: "Clinical Pathway" dashboard MIMIC-IV version 2

Figure 7.26 also includes general patient information in an information box, featuring the same variables and data as those used in the cluster analysis dashboard.

In addition, a combination of "Digital Twins" and "Clinical Pathway" is developed, like for the NAR database (Section 7.7.1).

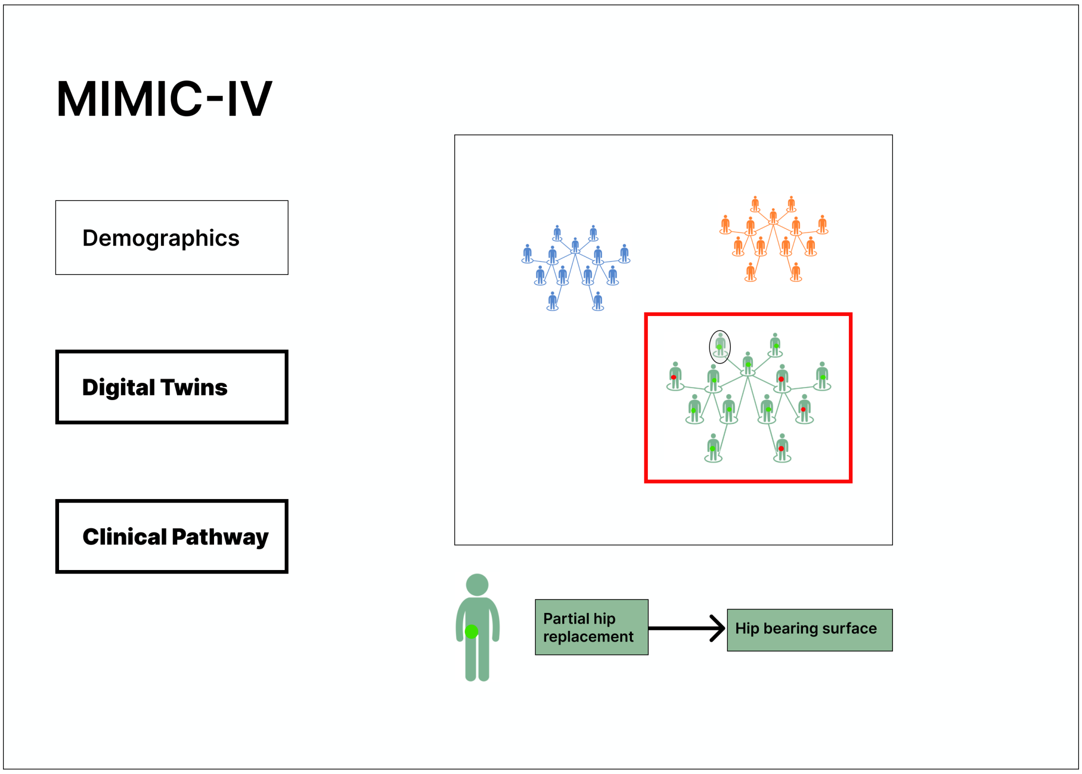

Figure 7.27: "Clinical Pathway" and "Digital Twins" dashboard MIMIC-IV

Walk-Through feedback on Fifth Iteration

The walk-through of the final design gave several suggestions for improvement. One suggestion to consider is to change the "Digital Twins" section to "Cluster Analysis" to better reflect the content, especially at this point. Adding titles for each section can enhance clarity, and moving forward, it would be beneficial to gather feedback from users to determine what additional information they need. A more detailed evaluation of the (final) design will be presented in the next chapter.

Chapter 8

Evaluation

In this chapter, a design evaluation is presented.

8.1 Evaluation

The concept of a digital twin in healthcare involves creating a virtual model of a patient, integrating various data sources to provide a comprehensive and dynamic representation of the state of the patient. Visualizations developed for this purpose must be intuitive, informative and aligned with the needs of healthcare professionals. Therefore, a thorough evaluation is essential to validate the design choices and their practical implications.
For the evaluation, I focused on revisiting the requirements established in Chapter 6, to determine if they have been fulfilled. Simultaneously, the focus has been on evaluating the usability and functionality of the design through several walk-throughs made throughout the iteration process and can be found in Chapter 7.

8.1.1 Requirements

Functional requirements

- **Present information about patient status.**
  In the example dashboard that presents Digital Twins and Clinical Pathways (Figures 7.18 and 7.27), the patients in the cluster are marked with a green or red mark to indicate their outcomes. This functionality should also be implemented in other dashboards representing digital twins / patients to consistently convey patient status.

- **Present information about the arthroplasty and its outcome.**
  Outcome information is visualized with a mark on the patients and further detailed through a clinical pathway. The visualizations of the clinical pathway in the NAR database (Figure 7.18) are related to a general clinical pathway in which each patient will go through some or all of the steps presented. In contrast, the MIMIC-IV database offers richer data, allowing the clinical pathway example to be more personalized. However, since the data are not yet fully clear and finalized, the visualizations currently represent generalized pathways and examples

rather than real data, as seen in Figure 7.27. Moving forward, outcomes details and additional information should be implemented, developed, and explored in greater depth, as this is very central to the overarching goal of the project.

- **Data analysis as a basis for digital twin definition.**
  The third requirement entails utilizing data analysis as the foundation for defining digital twins. In our project, cluster analysis has been the primary method of developing digital twins. However, the current results fall short of qualifying as digital twin representations and can be better described as digital cousins. At this stage, the visualizations conceptualize digital twins within clusters, acknowledging that accurate data implementation and visualization will be essential in the long run.

- **Produce understandable data visualizations.**
  The fourth requirement emphasizes the production of understandable data visualizations, a consideration that has been considered throughout the design process. The visualizations have been crafted to be simple and easily comprehensible, although there is a possibility that they may be perceived as overly simplistic, potentially lacking in additional information. It will be crucial to address this aspect further in subsequent development stages, especially as the project progresses and the data become more complex.

- **Include statistics whenever possible to provide healthcare personnel with details.**
  The fifth requirement, which involves including statistics whenever possible to provide healthcare personnel with details, has not been a top priority at the current stage. This is because the statistics and data complexity thus far have not been justified as an extensive presentation to the users. However, in the future, it will be crucial to gather more input from users, particularly healthcare personnel, to establish the information they deem important for implementation. This will help ensure that visualizations effectively meet the need and expectation of end users.

- **Differentiate visualization of patient-clusters and pathways.**
  The sixth requirement, which involves differentiating the visualization of patient clusters and pathways, has been addressed by implementing a distinct section on the dashboard for each aspect. The patient clusters are represented using color-coded visualization and are accompanied by information boxes (Figure 7.11). In contrast, pathways are represented using a combination of boxes, circles, and arrows (Figure 7.12). This clear differentiation ensures that users can easily discern between the two types of information presented on the dashboard. Another aspect is to present these elements in combination, which leads to the last functional requirement.

- **Integrate different elements of visualizations in a dashboard.**
  The seventh requirement focuses on integrating different elements of the visualization into a combined dashboard. This requirement has been approached more as a discretionary choice rather than as a complex visualization. Moving forward,

there is a need to develop a clearer strategy for how the combination of data elements will be implemented. This includes determining which elements of the two different parts of the visualization should be included to ensure coherence and usability.

Non-functional requirements

- **Aesthetics**
- **Flexibility**
- **Interoperability**
- **Testability**

Currently, the visualizations and prototype are quite simple and do not require any interaction, making the evaluation of flexibility and testability challenging. For aesthetics, color has been used to enhance visual appeal and comprehension of visualizations. Infographics and graphs have been incorporated to effectively display the data. Regarding interoperability, evaluation has been hindered by the lack of knowledge about the appearance and functionality of similar systems in the health care sector, making it difficult to determine how well the design integrates with other existing systems.
This evaluation highlights strengths and areas for improvement in the current visualization design, providing a foundation for further development. The following Chapters will explore these implications and outline the next steps to enhance visualizations and overall project outcomes.

Chapter 9

Discussion

In this Chapter, I am presenting a discussion of the visualizations, along with the diverse methods and methodologies employed throughout the research process. Furthermore, I address the research questions in detail.

9.1 Design Science Research

The design science research framework served as the overarching structure that guides the totality of this thesis. By integrating various methods into the research process, this framework, with its seven principles, provided an efficient and effective working environment. Design science research emerges as a highly recommended approach for domains that encompass real-world scenarios, necessitating the involvement of diverse expertise and methodologies.

9.1.1 Guidelines

Guideline 1: Design as an Artifact

In this thesis, I have developed a prototype dashboard that visualizes the outcomes and clinical pathways of digital twins. The artifact is a dashboard prototype that demonstrates the integration of solutions that are suitable for both databases and can serve as a cohesive visualization tool. This conceptual design is focused on visualization of a novel term, such as a digital twin, that relies on the results of data analysis and could accommodate various other data sources and analysis.

Guideline 2: Problem Relevance

The problem addressed is the need for an effective visualization of digital twin data in healthcare to improve understanding of patient outcomes and clinical pathways. By developing these visualizations, the research aims to provide healthcare professionals with better insights and decision-making support. The project explored how to apply this novel concept and make it useful in arthroplasty to contribute to outcome studies, patient safety, and generally better understanding of the patient population.

Guideline 3: Design Evaluation

The developed visualizations have been evaluated through cognitive walk-throughs with various experts working within the team or were associated with the team. These walk-throughs have assessed the effectiveness of visualizations in conveying the necessary information and supporting clinical decision-making. The walk-through was perceived as the best alternative for the evaluation of the visualizations, while the results that were included in the visualizations were assessed by the methods in terms of accuracy and significance.

Guideline 4: Research Contribution

This thesis has contributed by presenting a novel approach to visualizing digital twin data in healthcare (Chapter 7). It has also provided insight into the design process and methodologies used to integrate results from the analyses of the complex datasets, and implemented these results into a user-friendly dashboard prototype.

Guideline 5: Research Rigor

The construction of the visualization tool has been based on established design principles and data visualization techniques. Evaluation methods were applied to capture feedback from the project team at this point, but should involve in the future a wider set of users, such as healthcare professionals and biomedical researchers.

Guideline 6: Design as a Search Process

The design process has involved iterative development, where each iteration has refined the visualizations based on feedback and the results of the two other master students, Sahlgaard [30] and Røise [29], working with data mining. This iterative approach ensures that the final artifact effectively meets user needs and addresses the complexities of the datasets. System development methods, such as Kanban and weekly sprints, have contributed to gradual and iterative development due to the contribution of each team member.

Guideline 7: Communication of Research

The finding and development process is documented and thoroughly evaluated in the thesis. Additionally, the results are communicated through presentations and demonstrations in Chapter 7. This thesis will be open to the public at the University of Bergen Open Source Archive (https://bora.uib.no/bora-xmlui/).

9.2 Conceptual Design

Conceptual design has an advantage of showing how to design solutions without having all the data, results, or fully developed ideas regarding the future system or solution. In this project, the data and results came from two other parts of the project (Clustering [30] and Pathway [29] analyses) that ran in parallel with the visualization design.

This was possible since the concept for which visualization was created was a digital twin. My work has started with brainstorming about how to generate visualization and provide content that would suggest that this is a digital twin. The first question was what data analysis results to include and the second question is how many design solutions should be needed to present one twin. For example, should we have a presentation based on the cluster analysis only, another one based on the clinical pathways, and/or should we try to integrate all the results into one meaningful concept. These questions were also discussed with the team members, and my conclusion was that I will try to design for all three options. There was one more situation to consider, that is, two different databases, one was with the quality register and the other was with the MIMIC-IV database, which is an intensive care unit database. My thought was that visualization of the digital twin should not differ completely, even tho the databases are quite different in reality. However, the design should allow for the inclusion of specific results for different databases and allow the user to both orient themselves through the concept of digital twin paying attention to the specific sort of data. Another question that I thought about was how to integrate two aspects of the digital twin for the same database. My concern was that integrating both clusters and clinical pathways could be demanding if the data is very complex. My resolution was to keep one concept, but allowed navigation to specific results in which users could find details and look at the pathways, thus leading them freedom to interpret results. I thought it was of advantage to keep one visualization so that users have a possibility what kind of visualization they would like to explore. I assumed that in some situations they would prefer to just look at the results of the cluster analysis, while in other cases they would prefer to put that into the context of the clinical pathway.
Before Sahlgaard [30] and Røise [29] could come with their understanding of digital twins and pathways, I explored the nature and complexity of the databases at hand. This preliminary data analysis was performed to assess the quality and characteristics of the data. Descriptive statistics, data cleaning, and initial exploratory data analysis techniques were applied to understand the structure and patterns within the data. The analysis revealed significant challenges and limitations within the datasets that informed how I thought about visualization. For example, I found that descriptive statistics can result in interesting graphs that summarize data in a very efficient way, in other words, in terms of age, gender, and even county in the case of quality register data. This information can serve as some kind of identifying data that have both common sense and clinical meaning.

9.3 Visualization Development

The visualizations were designed and developed in 5 iterations (Chapter 7). The preliminary work on the literature has very few ideas about visualizing the digital twin. The most often was to see a flashy image in the context of theory or some schematic presentation, which has suggested the place for a visualized digital twin that could well represent and summarize common qualities of either patients or other subjects. On the other side, there were articles presenting models and processes in which there was discussion about digital twins. In spite of the great details presented in the models, it was hard to generate the concept of a digital twin. In my work I wanted to make a user-

friendly presentation that would clearly suggest how to find a digital twin and how a user can not lose its own way in the details. Given the nature of the data that was used in the project, I knew that it would be feasible to include the result of data analysis and connect them to the concept of digital twins. The resulting concepts were expected to offer clarity of navigation through the different outliers and finally obtain the particular data potentially seen as a digital twin to the current patient at hand. For example, if a surgeon needs to perform an arthroplasty of the hip for a 65-year-old female patient, he or she might want to know how things went for a digital twin. The conceptual design should allow this question to be answered in a simple as possible way, in a predictable manner, and leading to the concrete data to ensure that the physician has most patients similar to the one they are operating on.
In the current concept, there are four possible situations, one showing demographic data of the population, one belonging to the cluster analysis, one belonging to pathways, and one combining cluster analysis and clinical pathways. After a future evaluation with potential clinical users and researchers, these conceptual designs might undergo modifications and changes.

9.4 Limitations

Several limitations have affected this research. The databases used, the collaboration aspect, and the feedback given have had an impact on the final design and the process that led to it. Adjusting our initial expectations to a more complex clinical reality was necessary once we had reflected on the findings of both data mining and identifying clinical pathways.

The NAR Database

The NAR database is a quality register based on the patient electronic record which contains a large number of simplified patient records and lacks detailed information. This limitation hinders the depth of analysis and the ability to obtain comprehensive insights from the data.

The MIMIC-IV Database

The MIMIC-IV database presents complex data that is quite different from the NAR database, making the exploration and data mining more challenging. Unlike NAR, which consists of a single relational database, MIMIC-IV is a set of relational databases that can be integrated into one larger database. This structural difference made it more difficult to reuse the systems developed for the NAR database, which further complicates the integration process more than initially anticipated. Working with MIMIC-IV was a more time-consuming process.

Collaboration

Collaboration was an significant aspect of this thesis, involving three master's students in the project. Initially, the visualizations created were mainly conceptual, with the intention of integrating results from the other students later on. The results have shown that working with the data applying data mining generates many results, and not all of the data are eligible to be presented as a part of a digital twin. It is necessary to interpret the results, obtain the most significant and representative finding that reliably summarizes the patient population, and therefore contribute to the definition of digital twins. Similarly, not all pathways are simple straight forward and easy to include in the conceptual models, some patient cases show how complex their hospital stays are while the others go through the treatment in a simpler way but still with quite many clinical events which again shows diversity of pathways.

Feedback

The simplicity of the design and visualizations, which were less complex than originally anticipated, was reflected in the feedback. Feedback turned out to be less specific and limited to the concept. Given that the data and results of the analyzes were performed within the parts of the work dedicated to the cluster and pathway analysis. Therefore, I have taken a pragmatic view and focused primarily on conceptual design. For which I received feedback during walk-through sessions with the project members and those associated with the project. In the future, it will be crucial to obtain more information from end users, particularly healthcare personnel. These users can provide valuable input not only on the concept, but also regarding the additional requirements such as including information they need, excluding perhaps some of the features that are there now. It is to be expected that they will also have suggestions regarding the dashboard itself, and perhaps choice of color and fonts, and so on.

9.5 Answering the Research Questions

RQ1 - What is the best way to visualize the arthroplasty outcome analysis in the Digital Twin concept?

To effectively visualize the analysis of the results of the arthroplasty in the digital twin concept, clear color-coded indicators (e.g., Figures 7.18 and 7.27 were green for positive outcomes and red for negative outcomes) were essential. These indicators were integrated into the dashboard to show the current status of each patient's digital twin. In addition, detailed graphs and infographics were used to present relevant data, such as recovery times, complication rates, and overall success rates, providing a comprehensive view of patient outcomes. The use of infographics alongside statistical data improved both the aesthetic appeal and the understandability of the information.

RQ2 - What is the best way to visualize the arthroplasty clinical pathway of the Digital Twin concept?

To visualize the clinical pathway of arthroplasty within the digital twin concept, a step-by-step visual event log was used to detail each stage of the patient's journey. For the NAR database, this included preoperative, operative, and postoperative phases, marked with arrows to show the flow of the process. Color coding was used to differentiate between various stages and outcomes. For the MIMIC-IV database, clinical pathways are presented with a clear distinction between procedure and diagnosis, where procedures at this point can get a timestamp. Additionally, both frameworks facilitate the implementation of an interaction in which the user can click on each stage for more detailed information (such as procedure specifics, patient status, and expected recovery timelines), providing deeper insights and enhancing user engagement.

RQ3 - What are the alternatives for presenting the whole arthroplasty Digital Twin combining various analysis?

For this thesis, the dashboards were tested as the primary method for presenting the digital twin combined arthroplasty data. Dashboards are advantageous in providing a centralized real-time display of multiple data points with the opportunity to be customized and dynamically updated. They can be particularly useful for healthcare personnel who need a quick, comprehensive overview of patient data.
Preferably, it would be smart to explore different alternatives for presenting these data, including web-based designs. This exploration could provide further insights into the most effective ways to integrate and display the various elements of the digital twin concept, ensuring that the visualizations meet the diverse needs of end-users.

Chapter 10

Conclusions and Future Work

This Chapter concludes the thesis by summarizing the findings from the study and suggests future work. It also identifies issues that are still open to development and suggests possible directions for the future.

10.1 Conclusion

This master thesis aimed to develop and evaluate visualization methods for the digital twin concept in the context of arthroplasty outcomes and clinical pathways. The research focused on how best to represent patient status, the clinical pathway, and the integration of various analyses into a cohesive dashboard.
The project leveraged data from two databases: NAR and MIMIC-IV. The NAR database provided extensive patient data, but did not provide detailed information on surgeries and outcomes. In contrast, the MIMIC-IV database, though with fewer arthroplasty patients, offered richer clinical data, making it more complex to work with, but allowing for more detailed visualizations.
The key findings are the following:

1. **Visualization of Patient Outcomes:** Effective visualizations including clear, color-coded indicators to represent patient outcomes, enhancing the understandability of the data. The use of graphs and infographics further supports a comprehensive view of patient status.

2. **Clinical Pathway Visualization:** A step-by-step visual flow chart was effective in representing the clinical pathway of patients. Color coding and interactive elements provided a clear and engaging way to display each stage of the patient's treatment journey.

3. **Data Integration:** The use of dashboards proved to be a useful approach for integrating and displaying data from different analyses. However, the simplicity of the current visualizations suggests that further refinement and complexity is needed to fully meet the user's needs.

4. **Collaboration Challenges:** Working with two other master's students highlighted the importance of collaboration in developing a comprehensive tool. The integration of different contributions, particularly between front-end visualizations and back-end data mining, was more challenging than anticipated.

5. **Feedback and User Involvement:** Feedback from the walk-through indicated that although the visualizations were clear, they were perhaps too simplistic. It is a challenge to present lots of data and keep the visibility of data high, so a trade-off was made to include several features into the dashboard, but keep it as simple as possible. More detailed and specific feedback from end users, particularly healthcare professionals, will be crucial to future development.

10.2 Future Work

Building on the findings and limitations of this research, several areas for future work have been identified as follows:

Enhancing Data Detail and Accuracy

- Further refinement and completion of the dataset, particularly for the NAR database, to ensure accurate and comprehensive visualizations.
- Implementation of more detailed outcome data and information from the clinical pathway in the visualizations.

Exploring Alternative Visualization Methods

- Investigating the potential of web-based designs in addition to dashboards to provide more flexibility and accessibility.
- Developing interactive elements within the visualizations to allow for a more user-friendly experience.

Improving Collaboration and Integration

- Establish more efficient methods for integrating contributions from multiple team members, particularly in combining front-end and back-end developments.

User-Center Design and Feedback

- Engage healthcare personnel and other end users in the design and evaluation process to gather more detailed and actionable feedback.
- Conduct user studies to understand the specific needs and preferences of different user groups, such as general practitioners versus surgeons.

Interoperability and System integration

- Assess how the developed visualizations can be integrated with existing healthcare systems to ensure seamless data flow and usability.
- Developing standards and protocols for interoperability to facilitate the integration of different datasets and systems.

By addressing these areas, future research and development efforts can build on the foundation laid by this thesis, ultimately leading to more robust and effective visualization tools for the digital twin concept in healthcare. These advances will contribute to better patient outcomes and more efficient clinical workflows, aligning with the overarching goal of improving healthcare delivery through innovative technology.

Bibliography

[1] Cecilio Angulo et al. "A proposal to evolving towards digital twins in healthcare". In: *International work-conference on bioinformatics and biomedical engineering*. Springer. 2020, pp. 418–426.

[2] Patrizio Armeni et al. "Digital Twins in Healthcare: Is It the Beginning of a New Era of Evidence-Based Medicine? A Critical Review". In: *Journal of Personalized Medicine* 12.8 (2022).

[3] Maggie Mae Armstrong. "Cheat sheet: What is digital twin? Internet of things blog". In: *IBM* (2020). Accessed: 05/01/2024. URL: https://www.ibm.com/blog/iot-cheat-sheet-digital-twin/.

[4] Ankica Babic. *Digital Twin Project Documentation*. Project Documentation. Unpublished. 2023.

[5] Benjamin Bach et al. "Dashboard design patterns". In: *IEEE Transactions on Visualization and Computer Graphics* 29.1 (2022), pp. 342–352.

[6] Amanda Berlin. "What Are Event Logs and Why Do They Matter?" In: *Security Framework* (2023). Accessed: 01/05/2024. URL: https://www.blumira.com/what-are-event-logs-and-why-do-they-matter/.

[7] Bergthor Björnsson et al. "Digital twins to personalize medicine". In: *Genome medicine* 12 (2020).

[8] Canva. *Canva: Design, Art & AI Editor*. Accessed: 18/04/2024. URL: https://www.canva.com.

[9] Google Cloud. *BigQuery*. Accessed: 01/05/2024. URL: https://cloud.google.com/bigquery?hl=nb.

[10] David Concannon, Kobus Herbst, Ed Manley, et al. "Developing a data dashboard framework for population health surveillance: widening access to clinical trial findings". In: *JMIR formative research* 3.2 (2019).

[11] Angelo Croatti et al. "On the integration of agents and digital twins in healthcare". In: *Journal of Medical Systems* 44 (2020).

[12] Figma. *Figma: The Collaboratice Interface Design Tool*. Accessed: 18/04/2024. URL: https://www.figma.com.

[13] Git. *Git Guide*. Accessed: 01/05/2024. URL: https://github.com/git-guides.

[14] Hesledata. *Norwegian Arthroplasty Register (NAR)*. Accessed: 11/03/2024. URL: https://helsedata.no/en/forvaltere/bergen-hospital-trust/norwegian-arthroplasty-register/.

[15] Alan R Hevner et al. "Design science in information systems research". In: *Management Information Systems Quarterly* 28.1 (2004), pp. 75–105.

[16] Jane H Hodgkinson and Marc Elmouttie. "Cousins, siblings and twins: A review of the geological models place in the digital mine". In: *Resources* 9.3 (2020).

[17] IBM. *IBM SPSS software*. Accessed: 18/04/2024. URL: https://www.ibm.com/spss.

[18] A Johnson et al. *MIMIC-IV (version 1.0)*. 2020.

[19] Alistair E W Johnson et al. "MIMIC-III, a freely accessible critical care database". In: *Scientific data* 3.1 (2016), pp. 1–9.

[20] Peter Martinez. "What is Mid Fidelity Prototype and The Usages of It". In: *Mockitt* (2023). Accessed: 02/05/2024. URL: hhttps://mockitt.wondershare.com/prototyping/mid-fidelity-prototype.html.

[21] MCHP. "Concept: Charlson Comorbidity Index". In: (2023). Accessed: 02/05/2024. URL: http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?printer=Y&conceptID=1098.

[22] Johns Hopkins Medcine. *Arthroplasty*. Accessed: 02/03/2024. URL: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/arthroplasty.

[23] Johns Hopkins Medicine. *Hip Replacement Surgery*. Accessed: 02/03/2024. URL: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/hip-replacement-surgery.

[24] Andreas C. Müller and Sarah Guido. *Introduction to Machine Learning with Python*. O'Reilly Media Inc., 2017.

[25] Midori Nediger. "What is an Inforgraphic? Examples, Templates & Design Tips". In: *Venngage* (2023). Accessed: 12/02/2024. URL: https://venngage.com/blog/what-is-an-infographic/.

[26] Pandas. *Pandas*. Accessed: 18/04/2024. URL: https://pandas.pydata.org/.

[27] Jennifer Preece, Helen Sharp, and Yvonne Rogers. *Interaction design: beyond human-computer interaction*. 5th ed. John Wiley & Sons, 2019.

[28] Python. *Welcome to Python.org*. Accessed: 18/04/2024. URL: https://www.python.org.

[29] William Røise. *Leveraging Digital Twins for Clinical Pathways: Exploring Arthroplasty Registry and Clinical Database*. Submitted: 31/05/2024. (2024).

[30] Carl Oscar Kraft Sahlgaard. *Developing Digital Twins for Predictive Outcome Analysis in Arthroplasty*. Submitted: 31/05/2024. (2024).

[31] Scikit-learn. *Scikit-learn: Machine learning in Python*. Accessed: 18/04/2024. URL: https://www.scikit-learn.org.

[32] Wei Shengli. "Is human digital twin possible?" In: *Computer Methods and Programs in Biomedicine Update* 1 (2021).

[33] Ingrid Spilde. "Hvor mye bør du veie?" In: *forskning.no* (2019). Accessed: 02/05/2024. URL: https://www.forskning.no/overvekt/hvor-mye-bor-du-veie/1273135.

[34] Tianze Sun, Xiwang He, and Zhonghai Li. “Digital twin in healthcare: Recent updates and challenges”. In: *Digital Health* 9 (2023).

[35] Tianze Sun et al. “The Digital Twin in Medicine: A Key to the Future of Healthcare?” In: *Frontiers in Medicine* 9 (2022).

[36] Peyman Toreini et al. “Designing attentive information dashboards”. In: *Journal of the Association for Information Systems* 23.2 (2022), pp. 521–552.

[37] Takanori Yamashita et al. “Visualization of key factor relation in clinical pathway”. In: *Procedia Computer Science* 60 (2015), pp. 342–351.

[38] Min-Ren Yan, Lin-Ya Hong, and Kim Warren. “Integrated knowledge visualization and the enterprise digital twin system for supporting strategic management decision”. In: *Management Decision* 60.4 (2022).

[39] Mengdie Zhuang, David Concannon, and Ed Manley. “A framework for evaluating dashboards in healthcare”. In: *IEEE Transactions on Visualization and Computer Graphics* 28.4 (2022), pp. 1715–1731.

Appendix

A. Sikt Approval

Meldeskjema

Referansenummer
871912

Hvilke personopplysninger skal du behandle?

- Navn (også ved signatur/samtykke)
- Fødselsdato
- Lydopptak av personer
- Bakgrunnsopplysninger som vil kunne identifisere en person

Beskriv hvilke bakgrunnsopplysninger du skal behandle
Ønsker å behandle opplysning om personers navn, kjønn og yrke, sammen med alder (tall).

Prosjektinformasjon

Tittel
Visualization of Digital Twins

Sammendrag
"The Digital Twin project" aims to leverage the Norwegian Arthroplasty Register data to construct and utilize digital twins with the goal of improving predictions, recommending better therapies, optimizing treatment plans, as well as enabling preventive measures to combat points of failure. Digital twins are a digital representation of a physical person or object. This is distinct from a simulation in that a digital twin can perform several tests simultaneously. We are three students working on this project with different focus areas and main tasks. Two of the students will focus on the back-end programming and identifying and representing Twins in the dataset, while one will focus on the front-end and visualizations of this data, for medical personnel. To evaluate the artifacts created in this project we require expert knowledge and feedback. This specific project will focus on the visualizations and front-end programming.

Dersom personopplysningene skal behandles til andre formål enn behandlingen for dette prosjektet, beskriv hvilke
Nei

Hvorfor er det nødvendig å behandle personopplysningene?
Opplysningene som samles skal brukes til å få tilbakemeldinger på personers mening om prosjektets visualiseringer, og vurdering av brukeropplevelsen av prototypen som utvikles. Opplysningene som samles angår en persons alder, yrke, teknisk kunnskapsnivå og deres mening om visualiseringene og prototypen. Alle opplysninger som samles skal gjøres anonyme så ingen informasjon brukt i oppgaven kan lede tilbake til en intervjudeltaker.

Prosjektbeskrivelse

Project description.pdf

Ekstern finansiering
Ikke utfyllt
Type prosjekt
Studentprosjekt, masterstudium

Kontaktinformasjon, student
Ida Wergeland Sævareid, idawergeland2@hotmail.com, tlf: 94135970

Behandlingsansvar

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

Prosjektansvarlig
Ankica Babic, ankica.babic@uib.no, tlf: +4755589139

Skal behandlingsansvaret deles med andre institusjoner (felles behandlingsansvarlige)?
Nei

Utvalg 1

Beskriv utvalget
Eksperter innen prosjektets område. (Helsepersonell, IT-eksperter og forskere)

Beskriv hvordan rekruttering eller trekking av utvalget skjer
Rekruttering er gjort gjennom direkte henvendelser med telefon eller mail og presenterer hva prosjektet går ut på og om de finner det ønskelig å delta. Deltakere blir gitt et dokument (consent form) som presenterer prosjektet og ber om tillatelse om å behandle informasjon fra dem. Informasjonen samles gjennom et personlig intervju og med at de fyller ut et spørreskjema (System Usability Scale).

Alder
18 - 90

Hvilke personopplysninger skal du behandle for utvalg {{i}}? 1
- Navn (også ved signatur/samtykke)
- Fødselsdato
- Lydopptak av personer
- Bakgrunnsopplysninger som vil kunne identifisere en person

Hvordan samler du inn data fra utvalg 1?

Personlig intervju

Vedlegg

Intervjuguide.pdf

Grunnlag for å behandle alminnelige kategorier av personopplysninger
Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

Papirbasert spørreskjema

Vedlegg

System Usability Scale.pdf

Grunnlag for å behandle alminnelige kategorier av personopplysninger
Samtykke (Personvernforordningen art. 6 nr. 1 bokstav a)

Informasjon for utvalg 1

Informerer du utvalget om behandlingen av personopplysningene?
Ja

Hvordan informeres utvalget?
Skriftlig (papir eller elektronisk)

Informasjonsskriv

Samtykkeskjema.pdf

Tredjepersoner

Skal du behandle personopplysninger om tredjepersoner?
Nei

Dokumentasjon

Hvordan dokumenteres samtykkene?
- Manuelt (papir)

Hvordan kan samtykket trekkes tilbake?
En deltaker i prosjektet som ønsker å trekke tilbake sitt samtykke kan gjøre dette ved å ta kontakt gjennom telefon eller mail til Ida Wergeland Sævareid eller behandlingsansvarlig hos Universitetet i Bergen, Ankica Babic.

Hvordan kan de registrerte få innsyn, rettet eller slettet personopplysninger om seg selv?
En deltaker i prosjektet som har fått opplysninger registrert kan få innsyn, rettet eller slettet ved å ta kontakt gjennom telefon eller mail til Ida Wergeland Sævareid eller behandlingsansvarlig hos Universitetet i Bergen, Ankica Babic.

Totalt antall registrerte i prosjektet

1-99

Tillatelser

Skal du innhente følgende godkjenninger eller tillatelser for prosjektet?

Ikke utfyllt

Personverntiltak

Oppbevares personopplysningene atskilt fra øvrige data (koblingsnøkkel)?

Nei

Begrunn hvorfor personopplysningene oppbevares sammen med de øvrige opplysningene

Vi ønsker å bedre forstå profesjonell bakgrunn noe som kan påvirkes av alder, kjønn, erfaring og yrke.

Hvilke tekniske og fysiske tiltak sikrer personopplysningene?

- Personopplysningene anonymiseres fortløpende
- Flerfaktorautentisering
- Endringslogg
- Adgangslogg
- Adgangsbegrensning

Hvor behandles personopplysningene?

- Maskinvare tilhørende behandlingsansvarlig institusjon
- Mobile enheter tilhørende behandlingsansvarlig institusjon

Hvem behandler/har tilgang til personopplysningene?

- Prosjektansvarlig
- Student (studentprosjekt)

Tilgjengeliggjøres personopplysningene utenfor EU/EØS til en tredjestat eller internasjonal organisasjon?

Nei

Varighet

Prosjektperiode

01.10.2023 - 01.07.2024

Hva skjer med dataene ved prosjektslutt?

Data med personopplysninger oppbevares midlertidig til: 01.08.2024

Hva er formålet med den videre oppbevaringen av dataene?

Dokumentasjonshensyn

Vil de registrerte kunne identifiseres (direkte eller indirekte) i oppgave/avhandling/øvrige publikasjoner fra prosjektet?

Nei

Tilleggsopplysninger

B. Citi Program - Completion Report

COLLABORATIVE INSTITUTIONAL TRAINING INITIATIVE (CITI PROGRAM)

COMPLETION REPORT - PART 1 OF 2

COURSEWORK REQUIREMENTS*

* Scores on this Requirements Report (Part 1) reflect quiz completions at the time all requirements for the course were met. The Transcript Report (Part 2) lists more recent quiz scores, including those on optional (supplemental) course elements.

- **Name:** Ida Wergeland Sævareid (ID: 13061216)
- **Institution Affiliation:** Massachusetts Institute of Technology Affiliates (ID: 1912)
- **Institution Email:** yik008@uib.no
- **Institution Unit:** Information Science and Media Studies

- **Curriculum Group:** Human Research
- **Course Learner Group:** Data or Specimens Only Research
- **Stage:** Stage 1 - Basic Course

- **Record ID:** 61280536
- **Completion Date:** 15-Feb-2024
- **Expiration Date:** 15-Feb-2027
- **Minimum Passing:** 90
- **Reported Score*:** 92

REQUIRED AND ELECTIVE MODULES ONLY	DATE COMPLETED	SCORE
Belmont Report and Its Principles (ID: 1127)	15-Feb-2024	3/3 (100%)
History and Ethics of Human Subjects Research (ID: 498)	15-Feb-2024	5/5 (100%)
Basic Institutional Review Board (IRB) Regulations and Review Process (ID: 2)	15-Feb-2024	5/5 (100%)
Records-Based Research (ID: 5)	15-Feb-2024	4/4 (100%)
Genetic Research in Human Populations (ID: 6)	15-Feb-2024	4/5 (80%)
Populations in Research Requiring Additional Considerations and/or Protections (ID: 16680)	15-Feb-2024	4/5 (80%)
Research and HIPAA Privacy Protections (ID: 14)	15-Feb-2024	4/5 (80%)
Conflicts of Interest in Human Subjects Research (ID: 17464)	15-Feb-2024	5/5 (100%)
Massachusetts Institute of Technology (ID: 1290)	15-Feb-2024	No Quiz

For this Report to be valid, the learner identified above must have had a valid affiliation with the CITI Program subscribing institution identified above or have been a paid Independent Learner.

This document was generated on 15-Feb-2024. Verify at:
www.citiprogram.org/verify/?kaff59393-577c-4440-8b5b-f43742ce3943-61280536

Collaborative Institutional Training Initiative (CITI Program)
101 NE 3rd Avenue
Suite 320
Fort Lauderdale, FL 33301 US

Email: support@citiprogram.org
Phone: 888-529-5929
Web: https://www.citiprogram.org

COLLABORATIVE INSTITUTIONAL TRAINING INITIATIVE (CITI PROGRAM)
COMPLETION REPORT - PART 2 OF 2
COURSEWORK TRANSCRIPT**

** Scores on this Transcript Report (Part 2) reflect the most current quiz completions, including quizzes on optional (supplemental) elements of the course. The Requirements Report (Part 1) lists the reported scores at the time all requirements for the course were met.

- **Name:** Ida Wergeland Sævareid (ID: 13061216)
- **Institution Affiliation:** Massachusetts Institute of Technology Affiliates (ID: 1912)
- **Institution Email:** yik008@uib.no
- **Institution Unit:** Information Science and Media Studies

- **Curriculum Group:** Human Research
- **Course Learner Group:** Data or Specimens Only Research
- **Stage:** Stage 1 - Basic Course

- **Record ID:** 61280536
- **Current Score**:** 92

REQUIRED, ELECTIVE, AND SUPPLEMENTAL MODULES	MOST RECENT	SCORE
History and Ethics of Human Subjects Research (ID: 498)	15-Feb-2024	5/5 (100%)
Basic Institutional Review Board (IRB) Regulations and Review Process (ID: 2)	15-Feb-2024	5/5 (100%)
Belmont Report and Its Principles (ID: 1127)	15-Feb-2024	3/3 (100%)
Records-Based Research (ID: 5)	15-Feb-2024	4/4 (100%)
Genetic Research in Human Populations (ID: 6)	15-Feb-2024	4/5 (80%)
Populations in Research Requiring Additional Considerations and/or Protections (ID: 16680)	15-Feb-2024	4/5 (80%)
Research and HIPAA Privacy Protections (ID: 14)	15-Feb-2024	4/5 (80%)
Conflicts of Interest in Human Subjects Research (ID: 17464)	15-Feb-2024	5/5 (100%)
Massachusetts Institute of Technology (ID: 1290)	15-Feb-2024	No Quiz

For this Report to be valid, the learner identified above must have had a valid affiliation with the CITI Program subscribing institution identified above or have been a paid Independent Learner.

This document was generated on 15-Feb-2024. Verify at:
www.citiprogram.org/verify/?kaff59393-577c-4440-8b5b-f43742ce3943-61280536

Collaborative Institutional Training Initiative (CITI Program)
101 NE 3rd Avenue
Suite 320
Fort Lauderdale, FL 33301 US

Email: support@citiprogram.org
Phone: 888-529-5929
Web: https://www.citiprogram.org