Ingunn Skjevdal Hansen & Gry Sundt

PAsTAs - Visualization of Patient Trajectories

TDT4501 - Computer Science, Specialization Project, fall 2014

Software
Department of Computer and Information Science
Faculty of Information Technology, Mathematics and Electrical Engineering

Supervisors: Rune Sætre and Øystein Nytrø

NTNU
Norwegian University of
Science and Technology

ABSTRACT

With the large amounts of records of data stored on a patient, it can be strenuous for the patient to understand it all. Offering a visual illustration of the data can make it more applicable and manageable to grasp. In this paper, the focus group is patients who have suffered a stroke. An attempt to find the best way to illustrate the patient's contacts with the health sector is conducted, in the interest of giving a summarized presentation of the contacts. It will also be attempted to incorporate guidelines presenting possible ways to treat stroke patients into the illustration to the extent of giving more information on how the process is executed. Finally, a comparison of a single patient to a group of similar patients has been made in order to show how their treatment has been in comparison to their peers. By looking at the state-of-the-art and studying medical data, possible solutions for the illustration have been revealed. This paper is the first part of a Master Thesis at the Norwegian University of Science and Technology, which establishes the foundation for the timeline being developed in the second part of the project. The project is conducted under the supervision of the PAsTAs project.

CONTENTS

LIST OF FIGURES

ACRONYMS

ADL Activities of Daily Living

CSV Comma-separated values

Difi Agency for Public Management and eGovernment

EHR Electronic Health Records

GP General Practitioner

HELFO Norwegian Health Economics Administration

ICT Information and communications technology

IPLOS Individually based health and care service statistics

NST Norwegian Center of Integrated Care and Telemedicine

NTNU Norwegian University of Science and Technology

PAsTAs The Patients Trajectories Project

ACKNOWLEDGEMENTS

We would like to thank the members of The PAsTAs Project who allowed us to work on this project. Particulary Associate Professor Rune Sætre and Associate Professor Øystein Nytrø who have been the supervisors through the project. They are both employed at NTNU as well as being members in PAsTAs. Staff Engineer Håkon Wågbø has given valuable help in providing us with desired data from what PAsTAs has stored. Aslak Steinsbekk, Professor in Behavioural sciences in medicine and Health service research, is a member of PAsTAs and aided in the decision of the focus of the project, while Senior Researcher Gro Berntsen has given advice on how to group and identify the stroke debut of the patients. Finally, Coordinator Birgitte Mørch of the Norwegian Stroke Registry at St. Olavs Hospital has given guidance on how to interpret the patient trajectories.

1

INTRODUCTION

The health sector today is massive and complex. The different health divisions have their own types of procedures, and there is not sufficient communication between them. This is unfortunate, considering the number of patients who have interaction with multiple divisions throughout their treatments. Keeping a clear overview of what they are going through may be challenging for the patients if they do not get a summarized version of what they have been exposed to so far, and will encounter in the future. Some patients have regular communication with the health sector. One disease affecting a majority of people is chronic diseases. As much as 77 % of the disease burden in Europe is this kind of non-communicable disease (Nasjonalt folkehelseinstitutt [2010]). With comprehensive treatments and people living longer, the amount of medical data stored will become extensive. A tool to present an outline of all the medical data is needed in order for the patients to get a clear understanding of what they have been treated for and what the subsequent treatments will be.

1.1 BACKGROUND AND MOTIVATION

The Patients Trajectories Project (PAsTAs)[1] is a collaboration between the Norwegian University of Science and Technology (NTNU) and the Norwegian Center of Integrated Care and Telemedicine (NST), with the goal to develop methods for exploring data from electronic health records (EHR). Data stored in an EHR is an assembly of registered details about a patient in conjunction with getting some kind of health care[2]. Based on the data in the EHR, a *trajectory* illustrates the temporal[3] path of the patient. Showing patient trajectories to

[1] http://telemed.no/pastas-pasientforloep.5219575-247952.html
[2] http://www.kith.no/templates/kith_WebPage____569.aspx
[3] Relating to time

the patient it concerns has given PAsTAs the ability to evaluate the quality of the trajectory. Feedback from the patients after studying their own trajectory, has helped to compare the data registered in the EHRs to what the patient themselves perceived. PAsTAs supervised a master thesis the spring of 2014, in which this project has its basis (Wågbø [2014]). Since this project lasts until June 2015 and will consist of two different reports, there will be a differentiation between the *project*, which will refer to the whole project, and the *paper* which only refers to this text.

1.2 GOAL AND RESEARCH QUESTIONS

The purpose for this project is to make a visual representation of a patient's health records.

GOAL: *Present patient trajectories to give the patient a better overview on what services they have been in contact with in the health sector.*

The term *contact* refers to one visit with any part of the health sector, either a single point or an interval, in the EHRs. The data represented needs to be readable and understandable to people without a medical background, as well as having a simple enough visualization for people without a technical background. Patients should also see more details around their various contacts in order to get an improved understanding of what they have been through.

RESEARCH QUESTION 1 *What is the situation today for the visualization of temporal data?*

Looking at what is done previously provides a stable foundation when making decisions for the possibilities in the visualization system.

RESEARCH QUESTION 2 *What kind of information should a patient be able to see about themselves?*

There is an endless amount of data being stored when people are in contact with the different divisions of the health sector. The data would be too overwhelming for a patient to receive, and not all information is meant to be shared with patients. The question then becomes what data should be accessible for the patient, and how could the data be visualized?

RESEARCH QUESTION 3 *Is it possible to show (1) the individual patient trajectory, (2) the group of patients with the same illness, (3) and the clinical guidelines for the illness in a single visualization?*

The purpose of displaying a patient both their own trajectory as well as a distinctive group of patients is to show how an individual patient's treatment has been in comparison to patients with similar diseases.

PAsTAs conducted a workshop[4] where they wanted people to show how they would like their patient data data to be visualized. Out of this workshop, it became clear the participants would like to see the services they did *not* receive in a particular treatment, which should have been offered to the patient. By letting a patient compare themselves to a group with similar illness, it is possible to detect contacts they did not get themselves. If they see a majority has been through the same contacts as themselves, it might reassure them of receiving the best treatment for the severity of the stroke. By implementing the guidelines, the patient may see what they are expected to go through following the stroke. It provides patients with more reassurance the treatment received is according to the protocol. As well as getting the chance to see what they have been through, patients can see what the next steps will be. The options for the patient will be laid out, and help with understanding the information given by the doctor.

1.3 THESIS STRUCTURE

In order to obtain a overview of what this project is based upon, a thorough background of the medical and technical aspects will be presented. Secondly, the methods applied to extract more meaning will be described. With the use of the methods, some results will be explained. The results are discussed in the fifth chapter and finally, the research questions are revisited to determine the future tasks required to be conducted in the remainder of the project.

[4] http://www.ntnu.no/documents/21469517/1181072477/S%C3%A6tre%2C%20PAsTAs$_{H}elsIT.pdf$/296558$e2 - 112e - 43ea - b46a - d6dc50902c9f$

2

BACKGROUND

Visualizing patient trajectories is a technological task as much as a medical task. In order to visualize a trajectory and give valuable information to the patient it concerns, the data needs to be handled correctly. For this project, the scope of patients has been narrowed down to patients who have suffered a stroke, because it is a well-known medical condition that has been documented and researched extensively. The research has resulted in the development of guidelines which shows different routes stroke patients may follow. The technology available is also a thoroughly researched topic. Tools are developed regularly, and it is vital to know what exists to get a good judgment on what is a good solution to use.

2.1 STROKE GUIDELINES

Interpreting data of a stroke patient requires the understanding of the process of the treatment. All contacts a patient has which are related to the stroke will be called the *stroke treatment*. The Norwegian Directorate of Health has published guidelines for treating and rehabilitating stroke patients (Helsedirektoratet [2010]), Figure 1 shows a summarized figure of these guidelines. The guidelines illustrate how a stroke patient should be treated. It is a guide saying in what order the steps in the treatment should be executed, and what sequences should be avoided. For instance, a patient should always have an examination before being admitted to the stroke ward, as well getting an eventual admission into a rehabilitation facility should take place before getting a check-up at a polyclinic one to three months after the actual stroke happened. The initial stroke event will be referred to as the *stroke debut*. From looking at the guidelines, it becomes clear there are a number of different paths the patients can take. This is because there are many different levels of severity, and the patients need various kinds of assistance. Consequently, the stroke guideline must not be viewed as a standard trajectory for a stroke patient because of the variety from patient to patient.

Figure 1 explains the different possible paths a stroke patient goes through. The size of the arrows refers to how large a percentage of all stroke patients follow that line of procedures. All patients with symptoms of acute stroke should get to the emergency room ("Akuttmotak" in Figure 1) and then immediately be admitted to a hospital in the stroke ward ("Slagenhet"). Patients without symptoms should be checked by a doctor and based on this check-up either be admitted to the hospital (high risk) or be referred to get a more thorough examination ("Poliklinisk utredning") within 48 hours (low risk). After these two possibilities, there are a number of different routes the patient may end up following. If the stroke is too complicated, the patient may be transferred to the specialized stroke center ("Spesialisert slagsenter"), but all patients will get secondary prevention ("Sekundærforebygging") and early rehabilitation while they are admitted at the hospital. After the secondary prevention, the patient will be examined to decide the best course of action; home without rehabilitation, home with rehabilitation, nursing home, or admitted to a rehabilitation department. Patients should also get offered to come back for a checkup ("Poliklinikken") one to three months after discharge from the hospital.

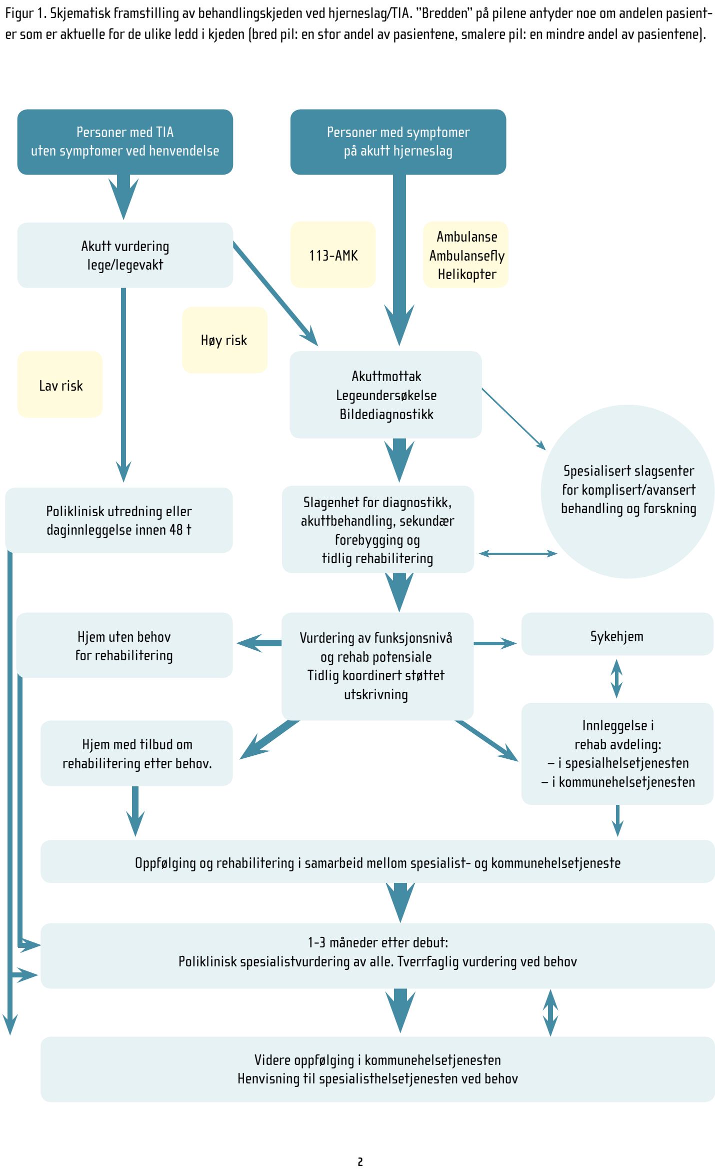


Figure 1: Stroke Guidelines

3

METHODS

This chapter describes the different methods being used throughout the project. Searching for relevant articles to obtain a precise picture of what is done so far in this field of study was the first priority. The decisions of what kind of tools that are going to be used are listed, and a mantra is introduced to aid in creating a user-friendly system. Finally, different aspects of the data are assessed.

3.1 RESEARCH METHOD

NTNU has their own search engine[1] to be able to find books and articles. By using key words such as "electronic health records" and "visualizing health records". Google Scholar[2] is also a search engine which has been of great use. Here, the same key words have been applied. The members of the PAsTAs project have throughout the project provided helpful advice, such as authors worth looking at for relevant articles. The project done by Wågbø [2014] has also been of help, by giving more relevant information, especially since the project was completed under the supervision of PAsTAs as well and had a similar topic. With the help of the recommended authors, articles of interest were found. By looking at the reference lists in the articles, more relevant articles were discovered as well as key words which could be used to find even more articles. This method of finding relevant information was used to find the answer to Research Question 1.

[1] http://www.ntnu.no/ub
[2] http://scholar.google.no/

3.2 GUIDELINES FOR INFORMATION VISUALIZATION

Shneiderman proposes the Visual Information Seeking Mantra for designing graphical user faces (Shneiderman [1996]). The mantra says how the visualization of data should be built up, "Overview first, zoom and filter, then details-on-demand". According to the mantra, users will first be met with an overview of the system, then by using tools available, be able to see more detailed information of the users' own choosing. The mantra is divided into seven tasks (overview, zoom, filter, details-on-demand, relate, history, and extract), which is encouraged to be used to interact with the data. These tasks are what constitute the base for what users can do with the interface.

The EHRs are made up of data, specifically time-oriented data. When the data is visualized, knowing what kind of time-oriented data is present is necessary. Aigner et al. [2008] identifies three distinctions. The first distinction was if the data had linear- or cyclic time. Linear time would mean the data has a starting point and a certain time domain. Cyclic time can for instance be the seasons where there is no need for any ordering of points since winter comes before summer, but succeeds summer as well. The second distinction is time points versus time intervals where time points have no duration and time intervals may last days or months. The final distinction is ordered time versus branching time. When the time is ordered, events happen one after the other. For branching time, there are multiple paths which gives different end scenarios. The article states the importance of choosing the correct visualization method for the data.

3.3 COLLECTION OF DATA

The data used in this project was provided by the PAsTAs project, and is stored in CSV files. It consists of EHRs from four different municipalities (Malvik, Melhus, Midtre Gauldal, and Trondheim), St. Olavs Hospital, and Norwegian Health Economics Administration (HELFO) over a two-year period. PAsTAs then identified the patients with chronic diagnoses, and requested to use their records in the project. Out of 16000 patients, 3000 agreed to answer a survey about their EHR. Since all patients were able to answer the survey, the data does not contain any patients who have died during the two-year period. Out of the 3000 patients, 117 had at least one contact with a stroke diagnose as the main diagnose in the dataset. When mentioning *dataset*, it is these 117 patients and their respective EHRs during the two-year period it refers to. The EHRs of the 117 patients will be what this paper will focus on. The information given in the dataset was:

- PID, patient identifier
- date, either one particular day (time point) or for a longer period (time interval)
- origin of data, either from a municipality, St. Olavs Hospital or HELFO
- unit- or service name; can specify a department in a hospital or specify a service given

All contacts a patient has had with the health sector are shown, which makes it difficult to know which contacts are connected to the stroke diagnosis. This problem was solved by marking contacts where stroke was the main diagnosis with the number 1, and the non-related contacts with the number 0. The marking of the number 1 will be called the *stroke diagnosis indicator*. Contacts represented with the number 1 will be referred to as being *stroke-related* throughout the paper. Appendix A lists the diagnostic codes used to identify a stroke, and these codes would make a contact stroke-related if stroke was the main diagnosis. Maintaining the anonymity of the patients was ensured by replacing patient names with a PID, as well as altering dates while still keeping the same temporal relations. With the new fictive dates, the data spanned from February 2010 to March 2012. The anonymity is also respected by not giving more information than what is needed. The dataset used in this project has been simplified, giving only the required information to present the various contacts. This means all the data is necessary for a patient to see, with regards to Research Question 2, which will be discussed further in the next chapters.

Identifying stroke debut

Getting an overview and a better understanding of the data is crucial when trying to find the best way to distribute information and visualizing it which is emphasized in Research Question 3. Making a prototype of how to visualize the patient data was initially done by making a paper-based timeline. Four patients with an average amount of contacts were selected, and their data was drawn together in the same timeline. The timeline is shown in Appendix B.1. The result of the paper-based timeline proved to be difficult to analyze and derive new information from. The patients had little in common, which concluded with the paper-based solution not being an optimal method. The next step was to have a more concentrated selection area for the patients. Patients with a contact at the stroke ward around the same week were grouped together. This is demonstrated in Appendix B.2. By looking at the result, it was clear the services and units unrelated to the stroke treatment polluted the picture. The different division who could have an affiliation with the stroke treatment was highlighted in pink. These contacts were presented using CHAP Links to get the timeline in a more clear and accurate format. All the patients represented in Figure 2 had a visit at the stroke ward at the end of May. Still, it is not apparent how these trajectories relate. Some of the patients had other diseases before the stroke which have blended with the stroke treatment. They may have abnormal stroke treatments, making them difficult to compare.

Pinpointing the stroke debut can be beneficial when looking for similarities in the patients' stroke treatment. The challenge with finding the debut, is the absence of an identifier to announce a stroke just happened. The stroke diagnosis indicator can only show the contacts connected to the stroke treatment. Also, since the dataset only has data for a period of two years, there is a possibility the stroke has happened before the two-year interval. A possible method to use to identify the stroke debut can be looking at patients who had their first stroke-related contact six months into the dataset. Most patients who suffer a stroke, have regular follow-ups with their GP, a physical therapist, or other services regularly throughout the dataset. It is therefore safe to assume the first stroke-related contact to appear the dataset will be the stroke debut when there has been silence for the last six months.

Grouping comparable patients

Another important subject in Research Question 3 is how to group patients together. The stroke patients' trajectories differ from each other. As seen in the stroke guidelines in Figure 1, there are multiple of ways to go through the process, and the amount of time spent at each location depends on the level of assistance needed by the patient. Some patients may have had a light stroke, and do not need any sort of rehabilitation, others might spend the next days or months in bed and need help with everyday tasks. It may therefore be relevant to group the patients to be able to see how the patients' trajectories resemble each other. This can also help visualize the possible treatment a specific patient can get by showing the treatment process of similar patients. Possible grouping criteria are:

- age
- function level
- previous medical history

Patients experience stroke differently when they are at the age of 40 compared to the age of 60. That is why this is a natural criteria for grouping patients together. At the St. Olavs Hospital in Trondheim, the patients' age affects what kind of hospital ward they end up in. If a patient is over 60, they are admitted to the stroke ward because they may need more assistance. The patients under 60 are placed in the neurology ward.

The severity of the stroke affects how well the patients will function later in their daily life. This is why patients are ranked on a scale depending on their degree of functionality, called *function level*. The identification of the function level is done differently depending on the origin of the data. The stroke ward and the neurology ward at St. Olavs Hospital use a scale called the Rankin Scale[3]. It ranks the patient on a scale from 0 to 6, where 0 is no symptoms and 6 is death. The municipalities have two different concepts to interpret the function level of a patient; IPLOS[4] and ADL[5]. IPLOS is a national registry used to describe the level of assistance a patient needs. By depicting a patients' assistance needs, it may give an image of how functional the patient is. ADL focuses on the daily activities a person performs from going to the bathroom to shopping groceries. ADL has 17 different variables and each one is measured on a scale of one to four (1-4). Health personnel measure a patient's ability to do these everyday activities, which shows how functional the patient is. By grouping patients based on their function level, they will most likely have similar processes when they are discharged from the hospital based services such as physical therapy and domestic assistance.

The third possible way to group patients is by looking at their medical history before the stroke. Some patients live a healthy life with little association to the health sector when the stroke suddenly takes place. Other patients may have continuous contacts with the health sector dealing with different kinds of hardships when the stroke takes place. Such patients may receive a different

3 http://www.strokecenter.org/wp-content/uploads/2011/08/modified_rankin.pdf
4 http://www.helsedirektoratet.no/kvalitet-planlegging/iplos-registeret/om-iplos/Sider/default.aspx
5 http://www.regjeringen.no/nb/dep/hod/dok/nouer/1997/nou-1997-17/7/5.html?id=345621

treatment. Patients getting treatment for an unrelated disease and suddenly suffer a stroke can either get treatment at their current ward, or get transferred to the stroke ward.

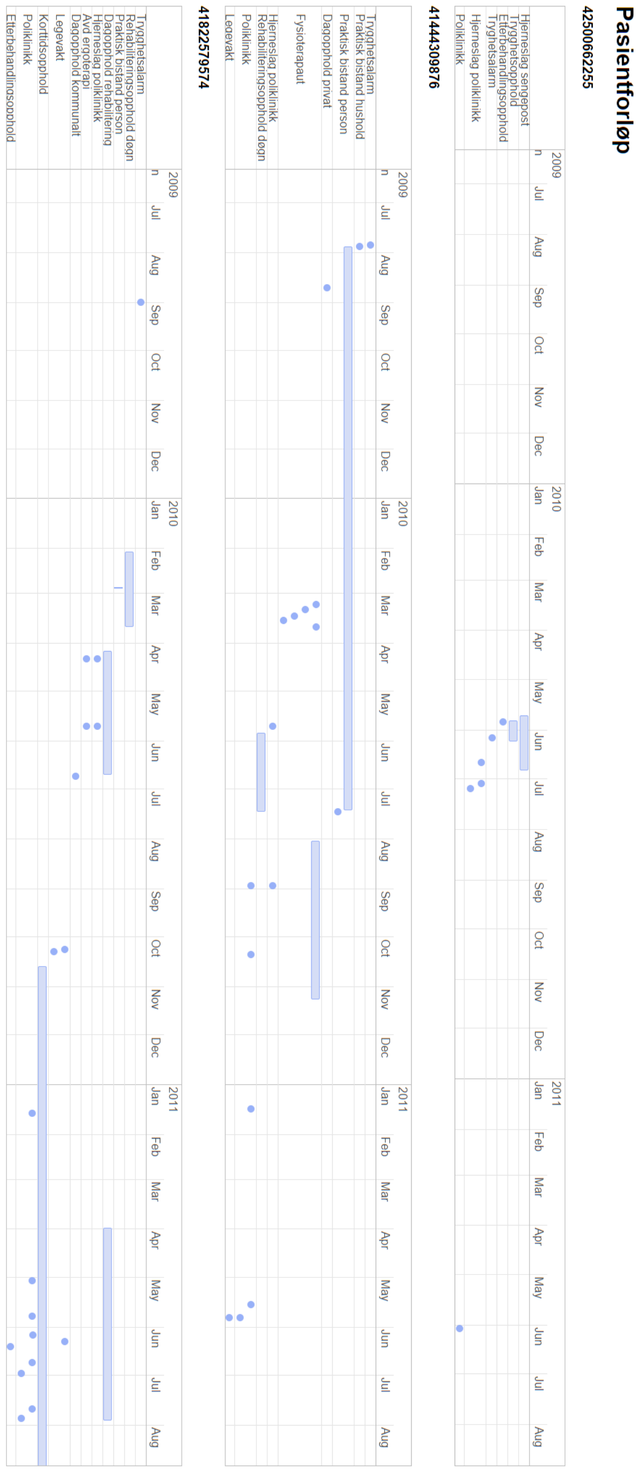


Figure 2: Timeline of patients with the same stroke debut time using CHAP Links.

4

RESULTS

Since this paper is a pilot project for the master thesis, there is more focus on gaining knowledge on the area of research which is connected to the project. The paper will result in giving a clear plan for what will happen further out in the project instead of a finished system. To understand and be able to interpret the data correctly for the different patients, a timeline was implemented. The timeline reflects the visualization criteria mentioned in Section 4.1. The method for finding similar patients is also described, and how the grouping of the patients can be performed. Finally, based on the research method described in the previous chapter, the state-of-the-art is reviewed.

4.1 STATE-OF-THE-ART

While EHRs are replacing paper-based health records, the need for systems to interpret the EHRs and help practitioners gain more knowledge, is growing. A human being cannot interpret the data without any kind of assistance. Based on the research method mentioned in the previous chapter, some articles gave valuable information for the state-of-the-art. Rind et al. [2010] lists up over 14 different systems developed to help clinicians make medical decisions based on the EHR data. The systems may also help clinicians gain new clinical knowledge by being able to ask the system questions and receive relevant patient data according to the parameters specified. Some properties reappear in multiple systems. A high level of interaction is present, where the user has various options in order to get more or less information on desiderata. Another property is the concept of alignment and the benefits it brings, is clear in the systems where it is present (Wang et al. [2008]; Plaisant et al. [2008]). The systems mentioned in the article all fulfilled the criteria listed below.

- The EHR consists of entire patient histories
- The information visualization deals with discrete non-spatial data
- The application is interactive with the user
- The system focuses on patient care, clinical research and quality control

Shahar [2013] also talks about how to aid in the process of interpreting and analyzing the time-stamped clinical data. The use of a mediator is proposed in order to be able to turn the raw data into more applicable, meaningful concepts. One of the systems handles multiple patients and the other focus on a single patient at a time. The systems concentrate on monitoring, visualizing, and exploring the data, which has given them good results in user tests.

The cancer ward at the St. Olavs Hospital in Trondheim has been a part of a pilot study with a system called eSP (Fremstad [2014]). This system shows a summarized, statistical overview of how the patients' treatments are being conducted in comparison to the standard guidelines. Such a comparison may increase the quality of treatment a patient receives, as well as make the work of the health personnel more efficient by calculating statistics and comparing the results against the standards. eSP was presented at the HelsIT conference[1], under the presentation "Pasientforløp og IT-støtte", and got good feedback from the audience. HelsIT was a two-day conference, consisting of 62 presentations focusing on the relationship between ICT and the health sector. All the participants were either from the health sector or people who worked closely with it. This section gives a summarized answer to Research Question 1, and will be further discussed in the next chapter.

4.2 INTERACTIVE VISUALIZATION CRITERIA FOR THE FINAL SYSTEM

In Section 4.1, the criteria for information visualization systems from the article by Rind et al. [2010] are listed. The final system being implemented in the string of 2015 will take into account the criteria. The first criteria stated the EHR consisted of the complete patient histories. Since the dataset provided by PAsTAs has a two-year scope, the criteria will not be realized. The system planned will make use of all available data, but focus on the stroke-related contacts. The system supports the second criteria, using discrete not-spatial data. There are no accurately referenced locations in the data, only hospital

[1] http://www.ntnu.no/helsit

wards and name of services. The third criteria mention interactivity. This is an essential part of the system. The patients, who will be the users of the system, will be able to interact with the system. Such interactions are zooming in and out to see different layers of detail. The last criteria do not fit into the planned system. All the systems mentioned was developed for the health care personnel, while the users in this system will be the patients.

4.3 TOOLS

After looking through numerous articles about how to visualize health records and seeing the importance of illustrating the temporal relations between different contacts, the decision resulted in making a timeline in this project. CHAP Links[2] was a tool that helped with constructing timelines. CHAP Links applies JavaScript and HTML5. Zotero[3] is a research tool aiding in collecting all references and effortlessly submit them into the paper. In order to collaborate on writing the paper, ShareLatex[4] has been used. It is a free, web-based editor tool helping with formatting and referencing among other valuable services.

4.4 THE STROKE DEBUT

Referring to Research Question 3 and the matter of comparing patients, an obstacle was pinpointing the stroke debut. The challenge with finding the patients' stroke debut was partially solved by using the method mentioned in Section 3.3. Using Microsoft Excel, the data was ordered by the date of each contact. Next, the patients who had a stroke diagnosis indicator during the first six months were filtered out. After this filtration, 79 out of 117 patients remained who all had their first stroke diagnose at some point between the fictive dates July 2010 and March 2012.

When looking over these patients manually, some patients resembled the stroke guidelines in Figure 1 more than others. The patients that resembled the most frequently taken path in the guidelines, were gathered into one group. By going through the selected 79 patients, the first stroke diagnosis indicator was analyzed to try to indicate if a stroke happened at this point. Such an indication

---

[2] http://almende.github.io/chap-links-library/
[3] https://www.zotero.org/
[4] https://www.sharelatex.com/

could be a patient visiting the emergency room or consulting their GP, and then the same day getting checked into the stroke center for a couple of days. Some patients had their first stroke diagnosis indicator too close to the end of the dataset, and were therefore excluded since it could not be correctly interpreted when too much of the data was outside of the time interval. All patients who did not have this typical start were excluded from the group of patients.

By these requirements, the list of 79 patients was narrowed down to 25 patients. Some of the patients' records were plotted into a timeline using CHAP Links, see Figure 3. In order to see any resemblance between the patients, it was necessary to align them according to a similar point in the trajectories. This point became the sentinel event of the stroke debut, presented with "Hjerneslag sengepost" in the figure. Alignment enhances the understanding and interpretation of temporal data Wang et al. [2008], which is important when it comes to patient trajectories and comparing the different treatments patients have gotten. In the timeline, all data available in the dataset is illustrated, because it is all relevant information giving a clear image of a patients' health.

The 38 patients initially excluded by removing patients with stroke diagnoses the first six months, was later analyzed using the same method as above. Five more patients fitted into the pattern with emergency room or GP, followed immediately by the stroke ward and were added to the group of initially 25 patients.

4.5 GROUPING PATIENT TRAJECTORIES

Another topic in Research Question 3 is how to group patients together. As suggested in Section 3.3, there are different ways to find similarities to group patients by. When taking into account what is useful, the solution is to group both age and function level. The function level will be deciphered by using the ADL data. Contemplating each variable of the ADL data, can give an insight into how well-functioning the patient is, and how much need for assistance there will be in the future. Patients around the same function level, may also have a comparable trajectory. The system will use the groups to structure the types of patients the records reflect. In the final product, patients will be able to compare their own trajectory to a group of related patients.

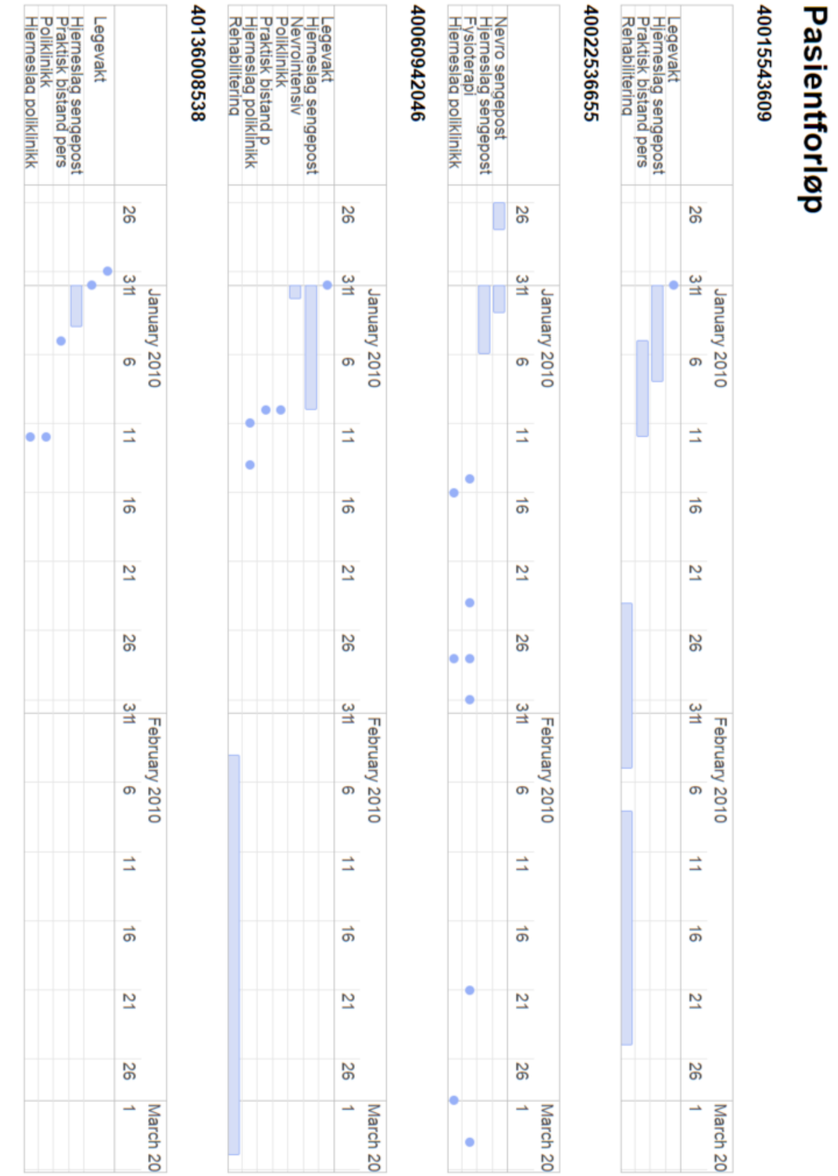


Figure 3: Trajectories of four patients, aligned at the stroke debut

5

EVALUATION

In this chapter, the topics mentioned previously are discussed and contemplated to find possible dangers or complications connected to them. Firstly, the state-of-the-art related to Research Question 1 are discussed. Secondly, the characteristics of the dataset are listed and what it implies for the project and Research Question 2. The remaining sections discuss different aspects connected to Research Question 3; the uncovering of stroke-related contacts and grouping patients.

5.1 STATE-OF-THE-ART EVALUATION

The similarity between the visualization systems mentioned in Section 4.1, is the focus on aiding the medical practitioners. There is little focus on visualizing the data in the patients' favor, which is the focus in this project in comparison to the systems existing today. The result from the master project conducted by Håkon Wågbø under the supervision of PAsTAs, is such a system making the EHR presentable for the patient it concerns. According to a survey conducted by Difi[1], six out of ten people in Norway would like to have electronic communication with their GP and access to their own patient trajectory online. To have access to a person's own patient record is a right protected by the rights act from 1999[2]. Many have taken on the challenge of developing systems to enhance the quality of treatment given to patients. This project will try to give the patients the system they need in order to be in control over their own health.

[1] http://www.difi.no/filearchive/digital-forstevalg-kartlegging-av-hindringer-og-muligheter-difi-rapport-2011-3.pdf
[2] https://lovdata.no/dokument/NL/lov/1999-07-02-63

5.2 CHARACTERISTICS OF THE DATASET

Based on the characteristics mentioned in Aigner et al. [2008], a better understanding of the project is given. The dataset can be defined as being in linear time (starts February 2010 with a time duration of two years), having both time points (contacts lasting no more than a day) and time intervals (contacts longer than a day), and the contacts are all in an ordered fashion. By looking at the various visual representations illustrated in the article and how each one handles the different characteristics of data, it becomes evident the choice of using a timeline to illustrate the patient trajectories is the best solution.

The stroke guidelines in Figure 1 is a kind of branching time, since there are several different routes the patients may go. Few visualization techniques exist for this kind of data (Aigner et al. [2008]). This could be an indication of the challenge to illustrate the stroke guidelines in the same timeline as the individual patients' which has a linear time. If it is not possible to incorporate the guidelines into the visualization for the patient, it could be a solution to have a simplified version of the guidelines beside the timeline. The reason is because the guidelines are a source of valuable information for the patient and should be incorporated in some way. If the patient can recognize what is stated in the guidelines and find similar contacts in their own timeline, the guidelines still give useful information.

5.3 STROKE DIAGNOSIS INDICATOR

Some uncertainties had to be considered when using the dataset. The data material identifies a specific contact as being stroke-related with the stroke diagnosis indicator, but it may be incorrect. For instance, the GP may have wondered if the patient had a stroke, and therefore wrote the diagnose down with a question mark. The patient could be taken to the stroke ward at the hospital for a more thorough examination. This examination may conclude the patient did not have a stroke, but the stroke diagnosis indicator from the GP will still be in the patients' records. The opposite scenario also occurs. There is no stroke diagnose, but the patient actually had a stroke or the contact was a part of the stroke treatment. One reason for this problem, is the stroke diagnosis indicator only takes considers the main diagnose. If a patient has more than one plausible or actual diagnosis, then the stroke can be missed by the filtering function. The diagnose ending up being listed as the main diagnose is not always the most critical. The result of this, is the stroke diagnosis indicator cannot be fully trusted to highlight the contacts with actual relevance to the stroke treatment.

Another uncertainty is patients not always ending up in the stroke ward or neurology ward after having a stroke. At St. Olavs Hospital, there is not always enough space at the ward the patient is supposed to be admitted to, and patients are instead placed in other wards with extra room. Not being sure whether or not the data reflects the truth makes it difficult to do a correct interpretation. A possible interpretation may be the patient was suspected of having a stroke, but after a false alarm was admitted for a different disease. Furthermore, a patient does have a stroke, and is admitted into a ward that is not stroke or neurology because those are full. An example may be two patients being admitted to the geriatrics ward with a stroke diagnosis indicator, but one of the patients is a false alarm. One patient may have been admitted to the geriatrics ward because the stroke ward was full, and the other patient was though to have a stroke, but instead was admitted to the geriatrics ward to get treatment for a different sickness. To detect such false alarms, it is necessary to look at the history as a whole, especially on what happens after the stroke debut, to observe if there are any indications of the patient getting treatment for a stroke. As mentioned in Section 2.1, stroke patients should be assessed at the polyclinic one to three months after the stroke debut. Looking for contacts like the stroke polyclinic can help determine if there was a false alarm or not.

5.4 identifying the stroke-related contacts

In this paper, the stroke patients have been compared by first finding the contact representing the stroke debut, and aligning the appropriate patients by this contact. Finding the stroke debut is only one way to conduct such a comparison. An approach like this could lead to patients being omitted as seen in Section 4.4 which may not be the desired result. Using another method may have the benefit of less patients being excluded. When one method is finding the starting point, a counter-method could be to find the ending point instead. Diseases and illnesses typically have a starting- and ending point. Finding the start- or end point should therefore be uncomplicated. The problem with this trail of thought is stroke normally causes longer periods of disabilities which may even last the rest of a patients' life, giving it a diffuse end point to the treatment. This is common for chronic diseases. Only having data over a period of two years also makes it difficult to find both the start- and the end point in this current dataset.

If the ending point is not a good approach, a possible choice is finding another specific contact in the timeline present for all patients. Fortunately, such a contact exists, which is the checkup at the polyclinic all patients should have one to three months after their stroke debut. Although it is not the end-point of the patients' treatment, it is a satisfying contact for comparison. The contact is clearly featured in the stroke guidelines as well, which can help with connecting the patients to the guideline.

Stated in research question 2, it is questioned whether or not it would be possible to illustrate one patient, multiple patients as well as clinical guidelines in one visualization. One of the reasons why the clinical guidelines should be presented, was to give the opportunity to compare the patients' treatment against what the regulations state the stroke patients are supposed receive of treatment. Here, looking at the end of the patients' EHR may be of value. If it is possible to identify where in the treatment process the patient is and transmit this spot into the guidelines, it could be possible to show the patient what steps in the treatment lie ahead. These steps can then be visually illustrated to the patient. Letting patients see their trajectory and a possible plan for their treatment after the stroke, may help them feel safer and not forgotten. It also gives them a better sense of control over their own situation.

5.5 GROUPING OF PATIENTS

For this project, it is decided to group patients based on their age and function level. A person at the age of 40 most likely has a different starting point than a person at the age of 60. This is emphasized by the division done at St. Olavs Hospital where patients under and over 60 years are admitted into different wards. Because of the lack of space for all stroke patients in one ward, the younger stroke patients are admitted to the neurology ward instead of the stroke ward. Through years of experience, it seems the hospital has detected the age of 60 as a good separator, while there is not room for all patients in the same ward. A group consisting of people the same age may give more relevant insight to a particular patient. Regardless of a patients' age, the severity of the stroke is always distinctive. That is why function level is a useful extra parameter to get groups of patients who have a more similar stroke treatment. If these two parameters would be used, it would result in specific groups of patients.

Since this project only consists of 117 patients, and the group has already been filtered down to 30 patients, it is important not to have a too high level of granularity. If the parameters are overly specific, it may result in groups with too few patients and not be of any use. If the number of patients would increase, it would be possible to make the parameters for the groups more distinguished and specific.

Unfortunately, few of the identifiers for functional level are present in the dataset. Without them, it is problematic to express how functional the patients are after their stroke. Some of the variables from IPLOS are used in the dataset, but there are other irregular variables of assistance as well. Not fully knowing what variables the timeline will consist of, makes it difficult to develop a system to handle all the data. If unexpected variables appear, exception-handling is needed to decipher them in order to depict precise function levels. Such exceptions can have a great impact on the system's efficiency and effectiveness Panzarasa et al. [2002]. A possible method for grouping patients further with the help of function level which would not require as much decoding would be dividing into two groups; the patients with no variants of assistance after discharge and the patients with assistance. Unfortunately, this method would result in few groups and give less similarities within them. A solution to this problem could be to incorporate the ADL-data into the dataset. If the ADL-data is more extensive than IPLOS and provides a better visual of the patients' function level, it may give a desired outcome.

6

CONCLUSION

The establishment of EHRs provided a helping hand to the health personnel when documenting medical data about patients. Subsequently, EHRs contain large quantities of interesting data which would normally be overwhelming to gather and analyze by hand. Storing the data electronically has opened up for a large amounts of possibilities on how to exploit the data; everything from aiding health personnel in making good clinical decisions to drawing parallel lines in various data to gain new medical knowledge. Countless systems have been developed in order to make it easier to manage the EHRs and aid the medical practitioners. But it is not only them who can benefit from this information. Patients have the right to get insight in their own health records, and such electronic records can make this task more manageable. In this paper, it is discussed options on how to visualize a patients' EHR in such a manner to make the temporal data apparent and gives valuable insight into their own health.

6.1 RESEARCH QUESTIONS

In order to evaluate the accomplishments of the paper, it is necessary to go back to the research questions. Since this paper is part of a larger project, not all questions are fully answered yet. Research Question 1 wants to know what characteristics the current state-of-the-art has and how this influences decisions for this paper. Multiple concepts have been taken into account based on the discoveries in the field of temporal information visualization. The results based on this question will be the pillars throughout the project.

Research Question 2 asks what kind of information a patient should see about themselves. So far, the data present in the dataset has been discussed, and mentioned that all current data should be visible to the patient, as well as the guidelines. However, this information is not enough. Further investigation on what information is accessible and which would be beneficial for the patients will be conducted.

Research Question 3 asks if it is possible to present multiple concepts in one visualization to provide the patient with more information. All three concepts have been discussed, with varying results as to how probable it is to incorporate into a single timeline. Illustrating one individual patient trajectory has been proven possible, different possibilities for grouping patients are mentioned, and the difficulties with the clinical guidelines are pointed out. The uncertainties will be looked at more thoroughly during the remainder of the project.

BIBLIOGRAPHY

Aigner, W., Miksch, S., Muller, W., Schumann, H., and Tominski, C. (2008). Visual methods for analyzing time-oriented data. *IEEE Transactions on Visualization and Computer Graphics*, 14(1):47–60.

Fremstad, J. (2014). St. Olavs Hospital - skreddersydd system for pasienten. `http://www.stolav.no/no/Nyheter/Skreddersydd-system-for-pasienten/130367/`. Accessed: 2014-12-07.

Helsedirektoratet (2010). Behandling og rehabilitering ved hjerneslag - kortversjon. pages 1–19.

Nasjonalt folkehelseinstitutt (2010). Folkehelserapport 2010 - helsetilstanden i norge. page 16.

Panzarasa, S., Madde, S., Quaglini, S., Pistarini, C., and Stefanelli, M. (2002). Evidence-based careflow management systems: the case of post-stroke rehabilitation. *Journal of biomedical informatics*, 35(2):123–139.

Plaisant, C., Lam, S., Shneiderman, B., Smith, M. S., Roseman, D., Marchand, G., Gillam, M., Feied, C., Handler, J., and Rappaport, H. (2008). Searching electronic health records for temporal patterns in patient histories: A case study with microsoft amalga. *AMIA Annual Symposium Proceedings*, 2008:601–605.

Rind, A., Wang, T. D., Aigner, W., Miksh, S., Wongsuphasawat, K., Plaisant, C., and Shneiderman, B. (2010). Interactive information visualization for exploring and querying electronic health records: A systematic review. Technical report, Technical Report HCIL-2010.

Shahar, Y. (2013). The elicitation, representation, application, and automated discovery of time-oriented declarative clinical knowledge. In Lenz, R., Miksch, S., Peleg, M., Reichert, M., Riaño, D., and Teije, A. t., editors, *Process Support and Knowledge Representation in Health Care*, number 7738 in Lecture Notes in Computer Science, pages 1–29. Springer Berlin Heidelberg.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pages 336–343. IEEE.

Wang, T. D., Plaisant, C., Quinn, A. J., Stanchak, R., Murphy, S., and Shneiderman, B. (2008). Aligning temporal data by sentinel events: Discovering patterns in electronic health records. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 457–466, New York, NY, USA. ACM.

Wågbø, H. D. (2014). The patient perspective: Utilizing information visualization to present health records to patients. Technical report, Norwegian University of Science and Technology, Department of Computer and Information Science.

Appendices

A

STROKE DIAGNOSES

Here is a list of diagnoses related to stroke, composed by PAsTAs. The contacts with the stroke diagnosis indicator will have one of the codes listed below as their main diagnosis. The codes are used at St. Olavs Hospital and HELFO.

Utvalg av diagnosekoder hjerneslag 210914

icd10	ICPC2	KronikerICPC2	DESCRIPTIO	ICPC2Txt
G460	K90	K	Arteria cerebri media-syndrom	Hjerneslag
G461	K90	K	Arteria cerebri anterior-syndrom	Hjerneslag
G462	K90	K	Arteria cerebri posterior-syndrom	Hjerneslag
G463	K90	K	Syndrom som skyldes hjernestammeslag	Hjerneslag
G464	K90	K	Syndrom som skyldes cerebellart slag	Hjerneslag
G465	K90	K	Rent motorisk lakunært syndrom	Hjerneslag
G466	K90	K	Rent sensorisk lakunært syndrom	Hjerneslag
G467	K90	K	Annet lakunært syndrom	Hjerneslag
G468	K90	K	Andre spesifiserte syndromer som skyldes hjernekarsykdom	Hjerneslag
I600	K90	K	Subaraknoidalblødning fra carotissifon og carotisbifurkatur	Hjerneslag
I601	K90	K	Subaraknoidalblødning fra arteria cerebri media	Hjerneslag
I602	K90	K	Subaraknoidalblødning fra arteria communicans anterior	Hjerneslag
I603	K90	K	Subaraknoidalblødning fra arteria communicans posterior	Hjerneslag
I604	K90	K	Subaraknoidalblødning fra arteria basilaris	Hjerneslag
I605	K90	K	Subaraknoidalblødning fra arteria vertebralis	Hjerneslag
I606	K90	K	Subaraknoidalblødning fra andre intrakranielle arterier	Hjerneslag
I607	K90	K	Subaraknoidalblødning fra uspesifisert intrakraniell arterie	Hjerneslag
I608	K90	K	Annen spesifisert subaraknoidalblødning	Hjerneslag
I609	K90	K	Uspesifisert subaraknoidalblødning	Hjerneslag
I610	K90	K	Subkortikal intracerebral blødning i hjernehalvdel	Hjerneslag
I611	K90	K	Kortikal intracerebral blødning i hjernehalvdel	Hjerneslag
I612	K90	K	Uspesifisert intracerebral blødning i hjernehalvdel	Hjerneslag
I613	K90	K	Hjernestammeblødning	Hjerneslag
I614	K90	K	Intracerebellar blødning	Hjerneslag
I615	K90	K	Intraventrikulær hjerneblødning	Hjerneslag
I616	K90	K	Hjerneblødning med flere lokalisasjoner	Hjerneslag
I618	K90	K	Annen spesifisert hjerneblødning	Hjerneslag
I619	K90	K	Uspesifisert hjerneblødning	Hjerneslag
I620	K90	K	Subduralblødning (akutt/ikke-traumatisk)	Hjerneslag
I621	K90	K	Ikke-traumatisk ekstraduralblødning	Hjerneslag
I629	K90	K	Uspesifisert intrakraniell blødning (ikke-traumatisk)	Hjerneslag
I630	K90	K	Hjerneinfarkt forårsaket av trombose i precerebrale arterier	Hjerneslag
I631	K90	K	Hjerneinfarkt forårsaket av emboli i precerebrale arterier	Hjerneslag
I632	K90	K	Hj.inf. forårs. av uspes. okkl. el. stenose, precereb. art.	Hjerneslag
I633	K90	K	Hjerneinfarkt forårsaket av trombose i hjernearterier	Hjerneslag
I634	K90	K	Hjerneinfarkt forårsaket av emboli i hjernearterier	Hjerneslag
I635	K90	K	Hj.inf. forårs. av uspes. okkl. eller stenose i hj.art.	Hjerneslag
I636	K90	K	Hj.inf. forårsaket av cerebral venøs trombose, ikke-pyogen	Hjerneslag
I638	K90	K	Annet spesifisert hjerneinfarkt	Hjerneslag
I639	K90	K	Uspesifisert hjerneinfarkt	Hjerneslag
I64	K90	K	Hjerneslag, ikke spesifisert som blødning eller infarkt	Hjerneslag

B

PAPER TIMELINES

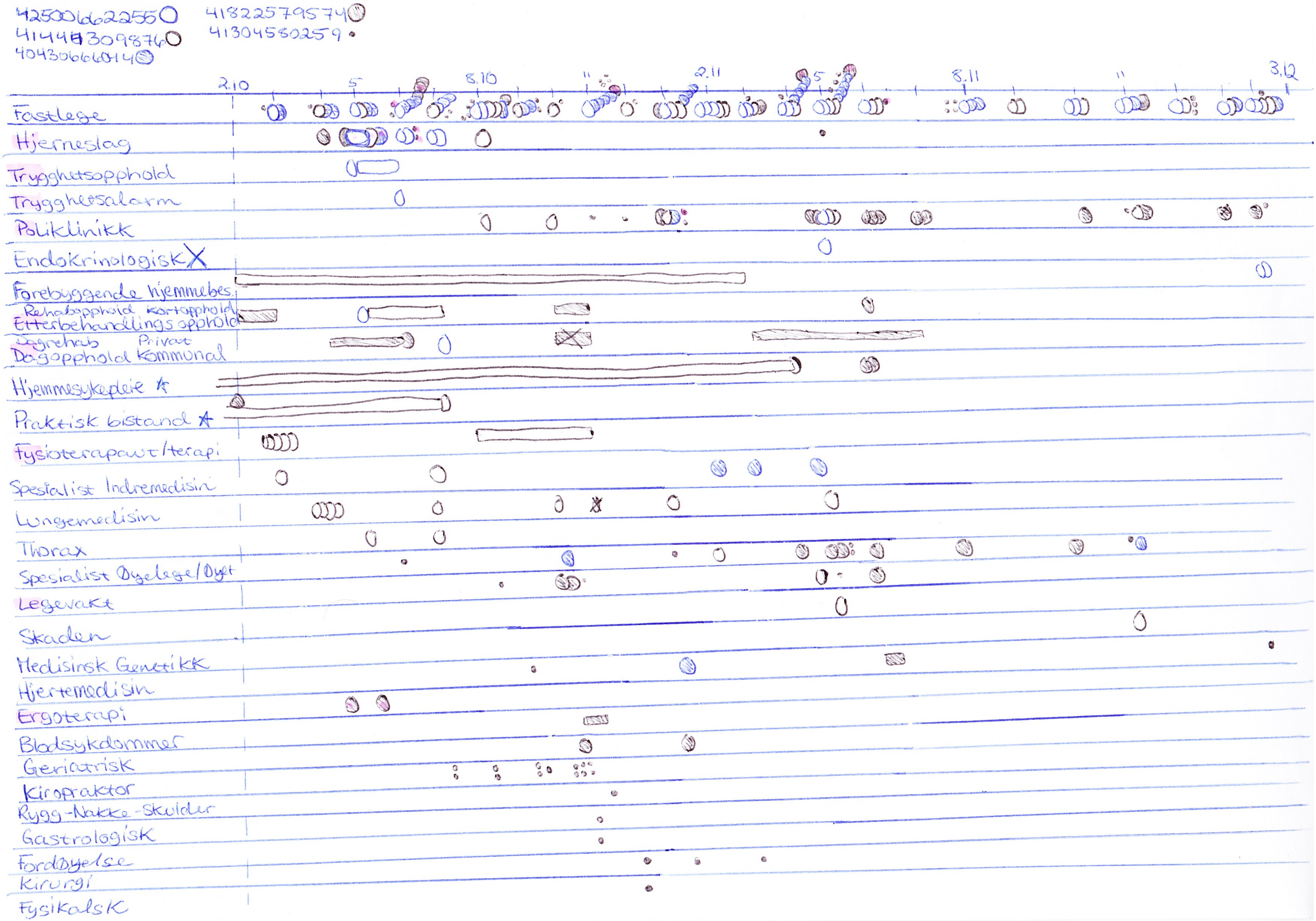
B.1 INITIAL TIMELINE

The first timeline was created to get a better understanding of the dataset, and different the patients could be. It was also made to investigate how it can look when more then one patient are inserted into the same diagram. The result of this is further discussed in Section 3.3.


P.ID: 41138687019
41822579574
40098203086
~~40356~~
40749083959
02.2010
5
08.2010
11
02.2011
5
08.2011
11
02.2012
Forebyggende Hjemmebesøk
Trondheim
Legevakt Kommunal
alt Helfo
Fastlege
alt Helfo
Øre Nese Hals Spesialist
alt Helfo
Spesialist
alt Helfo
Helsefremmende opphold
Trondheim
Legevakt
Helfo
Poliklinikk
Helfo
Øre Nese Hals
St. Olavs
Spesialist Hudlege
Helfo
Gastrologisk
St. Olavs
Hjerneslag
St. Olavs
Trygghetsalarm
Trondheim
Rehabopphold døgn
Praktisk bistand person
Ergoterapi
Blodsykdommer
St. Olavs
Geriatrisk
St. Olavs
Øyet
Helfo & St. Olavs
Hjertemedisin
St. Olavs
Hjemmesykepleie
Trondheim
Thorax
St. Olavs
Fastlønnet
Helfo

B.2 TIMELINE OF PATIENTS WITH THE SAME DEBUT TIME

After the initial timeline was analyzed, it was apparent that the patients in the same timeline, should have some common ground. They should also be aligned at some event. In this timeline the EHR of five patients which had their stroke debut within the same week was visualized.


42500662255
41449309876
40430666014
41822579574
41304580259
2.10
5
8.10
11
2.11
5
8.11
11
3.12
Fastlege
Hjerneslag
Trygghetsopphold
Trygghetsalarm
Poliklinikk
Endokrinologisk X
Forebyggende hjemmebes.
Rehabopphold
Kortopphold
Etterbehandlingsopphold
Dagrehab
Privat
Dagopphold Kommunal
Hjemmesykepleie
Praktisk bistand
Fysioterapaut/terapi
Spesialist Indremedisin
Lungemedisin
Thorax
Spesialist Øyelege/Øyet
Legevakt
Skaden
Medisinsk Genetikk
Hjertemedisin
Ergoterapi
Blodsykdommer
Geriatrisk
Kiropraktor
Rygg-Nakke-Skulder
Gastrologisk
Fordøyelse
Kirurgi
Fysikalsk