

# Knowledge-Assisted Visual Analytics: Data Exploration and Insight Generation of Health Care Data

DIPLOMARBEIT

zur Erlangung des akademischen Grades

**Diplom-Ingenieurin**

im Rahmen des Studiums

**Media and Human-Centered Computing**

eingereicht von

**Lisa Müllner, BSc.**

Matrikelnummer 01526345

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof. Mag.rer.soc.oec. Dr.rer.soc.oec. Silvia Miksch

Wien, 18. Mai 2021

---

Lisa Müllner

---

Silvia Miksch



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.



# Knowledge-Assisted Visual Analytics: Data Exploration and Insight Generation of Health Care Data

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

**Diplom-Ingenieurin**

in

**Media and Human-Centered Computing**

by

**Lisa Müllner, BSc.**

Registration Number 01526345

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Mag.rer.soc.oec. Dr.rer.soc.oec. Silvia Miksch

Vienna, 18<sup>th</sup> May, 2021

\_\_\_\_\_  
Lisa Müllner

\_\_\_\_\_  
Silvia Miksch



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Erklärung zur Verfassung der Arbeit

Lisa Müllner, BSc.

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 18. Mai 2021

---

Lisa Müllner



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Danksagung

Zuallererst möchte ich mich bei meiner Betreuerin Silvia Miksch für die tolle Unterstützung während des gesamten Prozesses dieser Arbeit bedanken. Außerdem möchte ich mich dafür bedanken, dass ich die Möglichkeit bekommen habe an diesem spannenden Thema zu arbeiten und genügend Freiraum hatte, um meine eigenen Ideen einzubringen. Weiters möchte ich mich bei Victor Schetinger, Velitchko Filipov und Davide Ceneda bedanken, die mich ebenfalls während des ganzen Prozesses meiner Arbeit durchgehend unterstützt haben und stets gute Ratschläge und wertvolles Feedback für mich hatten. Außerdem möchte ich mich auch bei der Fachexpertin Talin Barisani-Asenbauer dafür bedanken, dass sie uns den Datensatz zur Verfügung gestellt hat und auch stets mit fachlichem Wissen und Feedback zur Seite stand.

Ich möchte mich auch herzlich bei allen Teilnehmer\*innen der Evaluierungen dafür bedanken, dass sie sich die Zeit genommen haben und äußerst hilfreiches und detailliertes Feedback gegeben haben.

Weiters möchte ich meinen Eltern dafür danken, dass sie mich immer unterstützt und mir die Möglichkeit gegeben haben meine Ziele zu verfolgen. Genauso möchte ich auch meiner Schwester danken, die immer ein offenes Ohr und motivierende Worte für mich hatte!

Zum Schluss möchte ich noch meinem Freund Philip einen riesengroßen Dank dafür aussprechen, dass er mich immer unterstützt hat, meine Launen ausgehalten hat und auch die Ehre hatte, die gesamte Arbeit korrekturlesen.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.



# Acknowledgements

First of all I would like to thank my supervisor Silvia Miksch for the great support during the whole process of this work. I would also like to thank her for giving me the opportunity to work on this exciting topic and for giving me enough freedom to contribute my own ideas. Furthermore, I want to thank Victor Schetinger, Velitchko Filipov and Davide Ceneda, who supported me throughout the entire process of my work and always had good advice and valuable feedback for me. Another thanks goes to the domain expert Talin Barisani-Asenbauer for sharing the dataset with us and for always providing domain knowledge and feedback.

I would also like to thank everyone who took part in the evaluations for taking their time and providing extremely helpful and detailed feedback.

Beyond that I want to thank my parents who always supported me and gave me the opportunity to pursue my goals. Also a big thanks to my sister for always having an open ear and motivating words for me!

Finally, I would like to say a huge thank you to my boyfriend Philip, who has always supported me, had to endure my moods and also had the honor of proofreading the whole thesis.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Kurzfassung

Das Forschungsfeld “Knowledge-Assisted Visual Analytics” (KAVA) befasst sich mit der Integration von Domänenwissen in Visual Analytics Ansätze, die viele Vorteile für die Forschung sowie für die Analyse von Daten bieten, da sich Analytiker\*innen nicht auf ihr Domänenwissen verlassen müssen und sich mehr auf die Analyseaufgabe an sich konzentrieren können. Insbesondere im Gesundheitswesen verfügt KAVA über ein großes Potenzial, das derzeit nicht voll ausgeschöpft wird, da es nur wenige Ansätze gibt, die sich mit KAVA in Kombination mit Gesundheitsdaten befassen.

Um diese Lücke zu schließen, schlagen wir einen neuen KAVA-Ansatz vor, der die Möglichkeit bietet einen medizinischen Datensatz zu erkunden und zu analysieren. Der verwendete Datensatz ist aus einer klinischen Studie resultiert, bei der ein Medikament zur Behandlung der Augenkrankheit Uveitis geprüft wurde. Für den Entwurf und die Entwicklung des Ansatzes wird ein benutzerzentrierter Designprozess unter Einbeziehung einer Domänenexpertin in Kombination mit problemorientierter Visualisierungsforschung angewandt. Der finale Ansatz wird anhand einer qualitativen, aufgabenorientierten Nutzerstudie mit fünf Visualisierungsexperten validiert. Die Ergebnisse zeigen, dass der Ansatz in der Lage ist, sowohl die Analyse als auch die Exploration des Datensatzes zu unterstützen.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Abstract

The research area “Knowledge-Assisted Visual Analytics” (KAVA) deals with the integration of domain knowledge into Visual Analytics approaches which offers many advantages for research as well as for the analysis of data since analysts do not need to rely on their domain knowledge and can concentrate more on the analysis task itself. Especially in the health care sector, KAVA has great potential which is currently not fully exploited since there are only a few approaches that deal with KAVA in combination with health care data.

To fill this gap, we propose a new KAVA approach dealing with a dataset that resulted from a clinical trial of a medication for treating the eye disease Uveitis to provide the possibility of exploring and analyzing the dataset. For designing and developing the approach a user-centered design process, involving a domain expert, in combination with problem-driven visualization research is applied. The final approach is validated using a qualitative task-oriented user study with five visualization experts. The results suggest that the approach is able to support the analysis as well as exploration of the dataset.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Contents

<b>Kurzfassung</b>	<b>xi</b>
<b>Abstract</b>	<b>xiii</b>
<b>Contents</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Research Question and Sub-Questions . . . . .	2
1.3 Outline . . . . .	3
<b>2 Methodology</b>	<b>5</b>
2.1 Detailed literature research for different approaches . . . . .	5
2.2 Data exploration using the software Tableau . . . . .	5
2.3 Design and prototypical implementation of a new system . . . . .	6
2.4 Evaluation of the prototype . . . . .	7
2.5 Timeline . . . . .	8
<b>3 State of the Art</b>	<b>9</b>
3.1 Method . . . . .	9
3.2 Results . . . . .	10
3.3 Discussion . . . . .	18
<b>4 Data Exploration using Tableau</b>	<b>21</b>
4.1 Tableau . . . . .	21
4.2 Process . . . . .	22
4.3 Final Dashboards . . . . .	30
4.4 Data Exploration Findings . . . . .	33
<b>5 Design Process</b>	<b>39</b>
5.1 Problem Characterization . . . . .	39
5.2 Mockups . . . . .	44
<b>6 Prototype Implementation</b>	<b>57</b>
	xv

6.1	Technologies . . . . .	57
6.2	Graphical User Interface Composition . . . . .	58
6.3	Views and Components . . . . .	58
<b>7</b>	<b>Evaluation</b>	<b>69</b>
7.1	Preparation . . . . .	69
7.2	Evaluation Process . . . . .	70
7.3	Results . . . . .	71
<b>8</b>	<b>Conclusion &amp; Future Work</b>	<b>79</b>
8.1	Summary . . . . .	79
8.2	Future Work . . . . .	81
	<b>List of Figures</b>	<b>83</b>
	<b>List of Tables</b>	<b>89</b>
	<b>Bibliography</b>	<b>91</b>



# Introduction

## 1.1 Motivation

The health care sector is increasingly modernized and moving to more electronic solutions and technologies that store health care data, like the electronic health record (EHR) [CG15]. For research as well as analysis of health care data, this modernization provides many opportunities. However, due to the huge amount of data available, as well as its complexity, it is often hard to cope with the data, and the extensive potential is often not fully realized. Another example is data resulting from clinical trials where medical products are evaluated. Much of the data collected during those trials is available online, for example under [ClinicalTrials.gov](https://clinicaltrials.gov/)<sup>1</sup>. This data is also not widely used although the analysis of such data might lead to new insights that could help to improve health care and clinical research. The reason for this is again that the amount of data is huge and often only represented in a tabular form [Lam20].

Visual Analytics (VA) is one way of dealing with this problem by providing the possibility of visually exploring datasets with the help of interactive visualizations [KWEH14]. According to Thomas and Cook [TC06] Visual Analytics is defined as

“... the science of analytical reasoning facilitated by interactive visual interfaces.” [TC06, p. 10]

This is the reason why already many different approaches exist that visualize health care data. Especially for the visualization of patient data from the EHR, many approaches were already developed “to visually explore EHR data, gain insights, form hypotheses, and validate them” [FUA<sup>+</sup>15, p. 1].

---

<sup>1</sup><https://clinicaltrials.gov/> (accessed 01/05/2021)

However, one problem of the already existing solutions is that researchers and analysts need to rely on their domain knowledge when analyzing the data. This is the reason why the research area “Knowledge-Assisted Visual Analytics” (KAVA) came to life which deals with the integration of domain knowledge into Visual Analytic approaches to make analysis easier [FWR<sup>+</sup>17]. When integrating domain knowledge into visualizations analysts or researchers can concentrate more on the analysis task itself which is a big advantage since it could lead to even better results or new insights. Since KAVA is currently not very widespread the number of approaches utilizing domain knowledge in VA approaches in the health care sector is rather low [FWR<sup>+</sup>17].

Therefore, this thesis aims to fill this gap and provides a new VA approach dealing with KAVA. The thesis is conducted as a problem-driven visualization research project since it deals with a dataset received from a domain expert that resulted from a clinical trial where the efficacy of a certain medication for the treatment of the eye disease called “Uveitis” was tested.

### 1.2 Research Question and Sub-Questions

The main research question of this work is:

How can an interactive exploration environment supported by the integration of domain knowledge help and assist medical researchers to explore health care data of whole patient cohorts?

To narrow down the research field of the thesis the following subquestions have been added:

1. Which insights into a dataset and its structure can be gained when using a tool, like Tableau, to explore a dataset visually?
2. Which visualization techniques are most suitable for the visualization of health care data?
3. Which interaction techniques are necessary and suitable for the visualization of health care data?

Based on those questions the goal of this work is firstly, to provide an overview of the current state in the KAVA area with a focus on visualizing health care data, like patient data from the EHR or other sources, and secondly, to design, develop, and evaluate a prototype of an own approach that visualizes health care data. As already mentioned, the base for the prototype is a dataset from a clinical trial which is described in detail in Section 5.1.1. Furthermore, to get a better understanding of the data it will first be explored by using the software Tableau.

## 1.3 Outline

In Chapter 2 Methodology the methodology that was used for the thesis is presented. The results of the literature research are discussed in Chapter 3 State of the Art. After the literature research was finished the data exploration phase started and the process and results of this step are covered in Chapter 4 Data Exploration using Tableau. Based on the results and findings from the data exploration an own approach was designed whose process description can be found in Chapter 5 Design Process followed by the implementation of the prototype in Chapter 6 Prototype Implementation. The results of the evaluation of the prototype are discussed in Chapter 7 Evaluation. Conclusively, the last Chapter 8 Conclusion & Future Work summarizes the results of the thesis and gives an outlook to future work.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Methodology

In order to achieve the expected results, we used a methodology consisting of following four steps which are described in detail below:

1. Detailed literature research for different approaches
2. Data exploration using the software Tableau
3. Design and prototypical implementation of a new system
4. Evaluation of the prototype

## 2.1 Detailed literature research for different approaches

In the first step, we will search for literature in the field of “Knowledge-Assisted Visual Analytics” as well as in the more general area “Visualization of patient data”. We are expanding to the more general topic of patient data visualization, as there is a lack of systems in the “Knowledge-Assisted Visual Analytics” area due to missing research.

## 2.2 Data exploration using the software Tableau

The next step is then to explore the available data with the software Tableau. The goal of this step is to learn more about the data and to get a better understanding of datas’ structure and content. Furthermore, it should help to find out how to properly include the knowledge into the visualization.

### 2.3 Design and prototypical implementation of a new system

Based on the knowledge gained from the first two steps, the next step is the design and prototypical implementation of an own approach. For this step an user-centered design process in combination with problem-driven visualization research will be conducted. The problem-driven visualization research will be in the form of a design study as described by Sedlmair et al. [SMM12]. “A design study is a project in which visualization researchers analyze a specific real-world problem faced by domain experts, design a visualization system that supports solving this problem, validate the design, and reflect about lessons learned in order to refine visualization design guidelines” [SMM12, p. 2432].

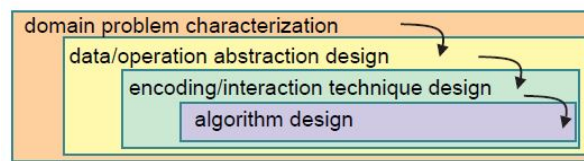


Figure 2.1: Nested Model [Mun09]: Base approach for the design process.

For the design of the visualization system two different approaches will be used. The first one is the nested model from Munzner [Mun09] and the second one is the design triangle from Miksch and Aigner [MA14]. The nested model consists of four nested layers (see Figure 2.1) that should help for the design and validation of the visualization. Therefore, in the first step we will define the data and the tasks that characterize the problem domain. This step overlaps with our second approach, the design triangle. The design triangle (see Figure 2.2) consists of three components which we will define for our approach.

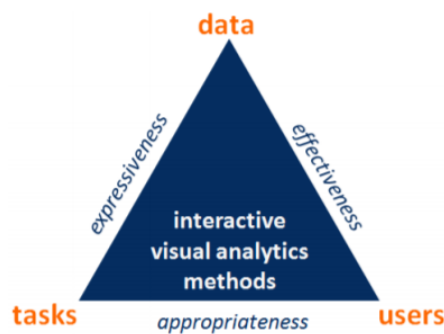


Figure 2.2: Design Triangle [MA14]: Components that we consider during our design process.

The first component is data which is in our case a dataset that resulted from a clinical trial of a medication that was conducted in multiple countries all over the world for the

	NEIVFQ25 VFQ1	NEIVFQ25 VFQ10	NEIVFQ25 VFQ11
VISIT	Overall Health	Noticing objects to side while walking	See how people react to things you say
Visit 1 (Baseline)	75	100	100
Visit 6 (Week 16)	50	100	50
Visit 8 (Week 24)	75	100	50
Visit 1 (Baseline)	50	25	25
Study Discontinuation	50	25	0
Visit 1 (Baseline)	50	75	100
Visit 6 (Week 16)	75	100	100
Visit 8 (Week 24)	75	50	100
Visit 1 (Baseline)	75	50	100
Study Discontinuation	75	50	100
Visit 1 (Baseline)	50	75	50
Visit 6 (Week 16)	25	50	50
Visit 8 (Week 24)	25	75	75

Figure 2.3: Snippet from the dataset that will be used as a base for the approach.

treatment of four different types of the eye disease Uveitis. The dataset includes general information regarding the patients, like gender, age, race, type of Uveitis, etc. During the study, the patients were asked to fill out two different questionnaires where both have the purpose of assessing the quality of life of the patients. Those questionnaires are called “Short-Form 36 questionnaire (SF36)” and “Visual Function Questionnaire (VF25)”. The patients answered those questionnaires three times during the whole trial. An example of a part of the anonymized questionnaire results can be seen in Figure 2.3. Furthermore, those two questionnaires have so-called subscores which are aggregated values of certain questions and those subscores are included in our dataset as well. However, a more detailed explanation of the dataset, the questionnaires, and the subscores is given in Section 5.1.1.

The second component are the users which are in this case mainly medical researchers and physicians. The last aspect are the tasks the user needs to perform which is in this approach the exploration of health care data.

After this first step is done we will proceed with the next layer of the nested model which is called “data/operation abstraction design”. The goal of this layer is to describe the data and operations defined in the first step in an abstract way using visualization vocabulary. Those descriptions are the input for the next layer where the visual encodings and interactions are designed. Finally, the last layer deals with the creation of an algorithm that implements the visual encodings and interactions.

## 2.4 Evaluation of the prototype

Finally, after the prototype is finished the prototype will be evaluated. For the evaluation a task-oriented qualitative user study with visualization experts will be conducted. The

study will be conducted with about 3-6 experts because studies have shown that this sample size is for an user study with experts enough [KP15; IIC<sup>+</sup>13].

The procedure of the user study will start by giving the experts a short introduction into the system and they are then asked to perform different tasks. The task-oriented approach will be enhanced with a thinking-aloud study to get also detailed observations of the experts themselves while performing the task [Lew82].

## 2.5 Timeline

In Figure 2.4 a schematic timeline of the whole thesis is illustrated that shows the temporal order of the main milestones as well as their main sub-tasks in between.

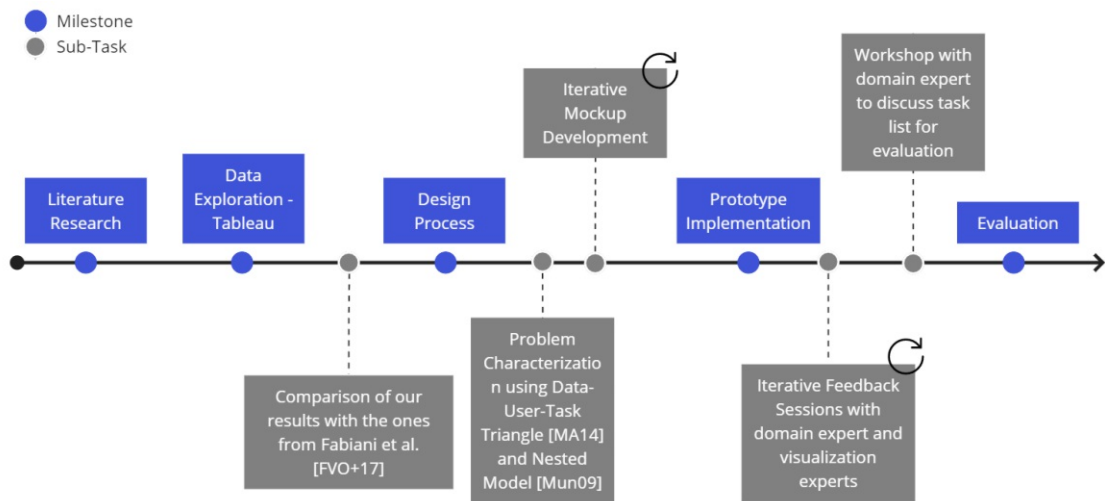


Figure 2.4: Schematic timeline of the thesis: All main milestones (blue color) are arranged in their temporal order as well as their main sub-tasks (grey color) that are conducted in between.



# State of the Art

The first step of the thesis was to research already existing approaches in the area of KAVA. The results of the research are presented in this chapter by providing an overview of the current state of the art of KAVA in the health care area.

First, the method for acquiring the approaches is shortly discussed, followed by the presentation of the found approaches. In the end, the approaches are compared and analyzed regarding their advantages and disadvantages.

## 3.1 Method

Before we started with the research of approaches in the KAVA area we first decided on our inclusion and exclusion criteria for the approaches we find. Since our interest is mainly on visualization of patient data in combination with domain knowledge, regardless of whether it is patient data from the EHR or other sources, like clinical trials, etc. the first criterion was to search for such approaches. Another criterion was to include only approaches that visualize data from whole patient cohorts since our approach is focusing on data from a larger cohort. Furthermore, one strict exclusion criterion we defined is that no commercial systems will be used since scientific literature is missing most of the time.

After those criteria were defined, we started our research by using the search engine Google Scholar<sup>1</sup>. We mainly used the following phrases and keywords as an input for the search:

- knowledge assisted visual analytics
- knowledge assisted visual analytics health care data

---

<sup>1</sup><https://scholar.google.com/> (accessed 04/05/2021)

- patient data visualization
- medical data visualization

Furthermore, we wanted to focus mainly on papers that are rather new and, therefore, filtered in Google Scholar for papers that were published since 2017. In case the filter did not lead to enough search results we switched the filter off again.

It did not take much time until we found the first approaches that fit perfectly into our criteria as well as the research field, which are the approaches from Federico et al. [FUA<sup>+</sup>15] as well as Gschwandtner et al. [GAK<sup>+</sup>11]. After those approaches were found we additionally searched for papers that cited the aforementioned papers as well as their references.

For all of the papers we found that seemed to fit our criteria we had the same procedure which was to read the abstract first and if the abstract sounded relevant to our topic we scanned the paper in more detail and decided whether to save it for later or not. In the end, we managed to find around 20 papers as candidates for our state of the art report and after reading all of them in detail we narrowed it down to seven approaches. All these approaches are presented and discussed in the next section.

## 3.2 Results

As already mentioned before, the number of approaches dealing with KAVA in the medical domain is rather low which is the reason why we searched not only for approaches that deal with KAVA and instead widened our search area to the more general topic of patient data visualization. The following section is, therefore, structured by the different types of approaches depending on whether they deal with KAVA or visualize only patient data. We start with KAVA-related approaches and present afterward more general patient data visualization approaches.

### 3.2.1 KAVA-related approaches

In the category of approaches dealing with the research area KAVA four approaches were found.

We start by presenting the oldest approach which is **CareVis** [AM06]. CareVis is an approach visualizing patient data in combination with domain knowledge represented as treatment plans, also known as clinical practice guidelines. CareVis presents, on the one hand, the structure of the guidelines, and, on the other hand, which actions of the plan were executed as well as their temporal order combined with visualizations of the patient data that is affected by the executed actions. This makes it possible to analyze whether the executed actions from the treatment plans were effective in improving the health of the patient or not.

CareVis' graphical user interface is composed of three views which are a logical view, a temporal view, and a QuickView panel as can be seen in Figure 3.1. On top of the screen, the QuickView panel is presented which has the goal of informing the user about the most important values of the current patient. Below the QuickView panel, the logical view, as well as the temporal view, are aligned next to each other. In the logical view, the treatment plans' structure is presented with a flowchart-like representation called AsbruFlow. The third view, the temporal view, visualizes, on the one hand, what actions of the treatment plan were executed at which time using the PlanningLines visualization [AMTB05]. On the other hand, the parameters that are intended to be affected by the executed action of the plan are visualized.

CareVis was evaluated by five physicians using a four-step process, that consisted of a short introduction, followed by the presentation of the prototype as well as the testing of it, and was closed with a feedback session including a questionnaire [AM06]. According to Aigner and Miksch [AM06] the physicians' responses were very positive emphasizing the well-designed structure as well as the intuitive graphical representations and symbols.

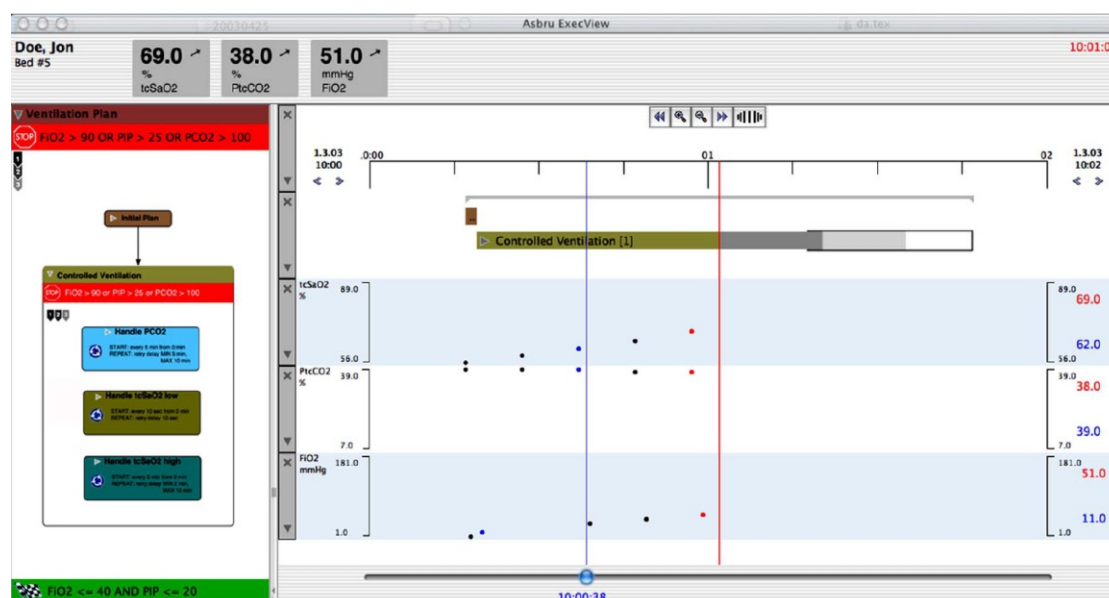


Figure 3.1: CareVis [AM06]: Visualization of treatment plans combined with patient data from a single patient. On the left, the structure of the treatment plan is visualized. Next to it, the executed actions of the plan are visualized as well as certain parameters the plan affected and on top is the QuickView panel providing information regarding the most important patient parameters.

The second approach in this category is called **CareCruiser** [GAK<sup>+</sup>11]. CareCruiser is a derivative of CareVis with the difference of visualizing multiple patients at once [GAK<sup>+</sup>11]. For each patient, the executed actions of the treatment plan are visualized as well as

### 3. STATE OF THE ART

important patient parameters and the patient records are then stacked above each other as can be seen in Figure 3.2. Furthermore, CareCruiser enhances the patient parameter visualization by using color-coding to see more easily how the actions affected the parameters. Another difference to CareVis is that CareCruiser does not use PlanningLines for the visualization of the treatment plans. Instead, diamonds are used to visualize the executed actions. To visually illustrate what type of action was executed Gschwandtner et al. [GAK<sup>+</sup>11] use differently colored diamonds for the different actions.

CareCruiser was evaluated using two different methods. First, they received feedback from a medical expert with whom they collaborated throughout the entire project. The second method was a heuristic usability evaluation that was conducted with four participants. During the heuristic evaluation 32 usability problems were found but according to Gschwandtner et al. [GAK<sup>+</sup>11] most of them “can easily be fixed, for instance, by adding tooltip text to buttons” [GAK<sup>+</sup>11, p. 48].

From the medical expert, some domain knowledge was provided, for example, that in the medical area a distinction is made between parameter values that are below or above an intended value and, therefore, those values should not have the same color.

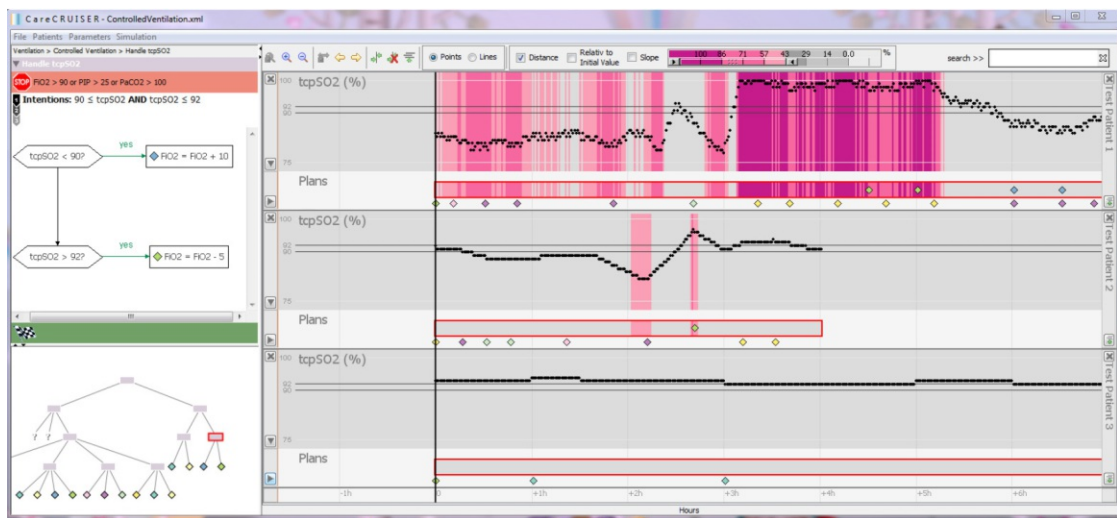


Figure 3.2: CareCruiser [GAK<sup>+</sup>11]: Visualization of treatment plans and multiple patient records. On the left, the treatment plan is visualized and on the right, the patient parameters are presented as well as the executed actions using diamonds.

Another approach that was found during research is **Gnaeus** [FUA<sup>+</sup>15]. The general structure of Gnaeus is similar to the one of CareVis and CareCruiser as can be seen in Figure 3.3. For the treatment plans the same visualization is used as well. The main difference between Gnaeus and CareVis/CareCruiser is the main view since Gnaeus visualizes data from patient cohorts using different visualizations.

Depending on the number of patients, two different visualizations are used for the

visualization of the patient parameters. For patient cohorts with a low number of patients Qualizon Graphs [FHR<sup>+</sup>14] are used for the visualization of the data of every single patient. If the cohort is larger more space-efficient visualizations are used which are depicted in Figure 3.3. As can be seen, the main view is split into three areas that are stacked above each other. The first one on top includes multiple streaming box plots. Those box plots are visualizing the aggregated values of the patient parameters. The second area shows a pixel-based visualization technique where each patient is represented by a row. For the pixel-based visualization, a red and blue color-coding is used where in the case of the data in Figure 3.3 a cell is colored blue if the value is lower than 80% and red otherwise.

The third and last area of Gnaeus' main view is responsible for visualizing the aggregated data regarding the executed actions from the treatment plan. Circles are used as representation of the actions whose alpha value encodes how often the action was executed within a time interval. Furthermore, the circle radius indicates how many patients got the action administered.

During the whole design process and development phase they collaborated with three experts. Additionally, a user study with two medical experts was conducted. According to Federico et al. [FUA<sup>+</sup>15] the response was very positive and the experts mentioned that they could imagine using Gnaeus in their daily routine.

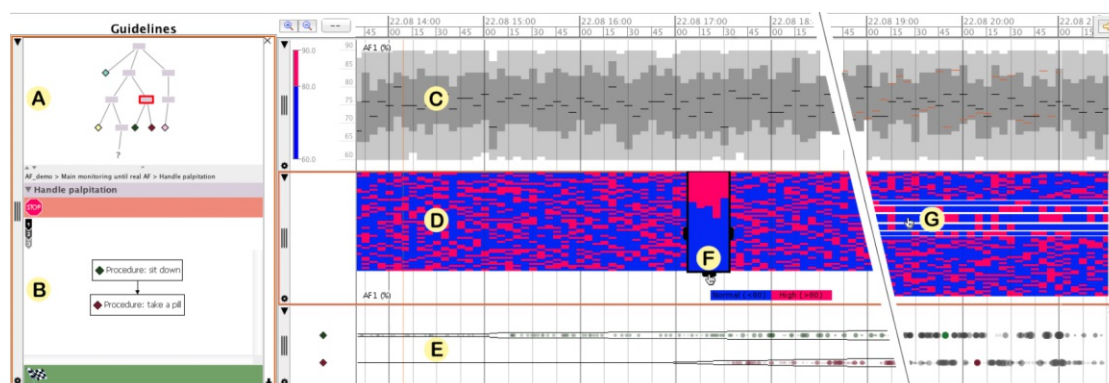


Figure 3.3: Gnaeus [FUA<sup>+</sup>15]: Visualization of the treatment plan structure and patient data of whole cohorts. (A) shows the hierarchical structure of the treatment plan whereas (B) displays the structure of the selected subplan. (C) is a streaming box-plot visualization of the aggregated parameter data for the selected subplan and (D) a pixel-based visualization of the relevant patient parameter for each patient. (E) depicts the visualization of the aggregated data regarding the executed actions and (F) the grouping lens interaction technique that visualizes the distribution as stacked bars. (G) shows the vertical fish-eye distortion available for detailed patient analysis.

The approach from Basole et al. [BPG<sup>+</sup>15] is the last one in this category. They propose a **web-based visualization** “to explore care processes and their conformance to guidelines”

[BPG<sup>+</sup>15, p. 2]. Their approach consists of a “Careflow” visualization, a “Conformance” visualization, and visualizations of statistical data from different patient cohorts.

For the “Careflow” visualization a Sankey diagram is used as can be seen in Figure 3.4. With the Sankey diagram the flows between executed activities, like lab test results or medications are visualized which makes it possible to find out which paths are more followed than others.

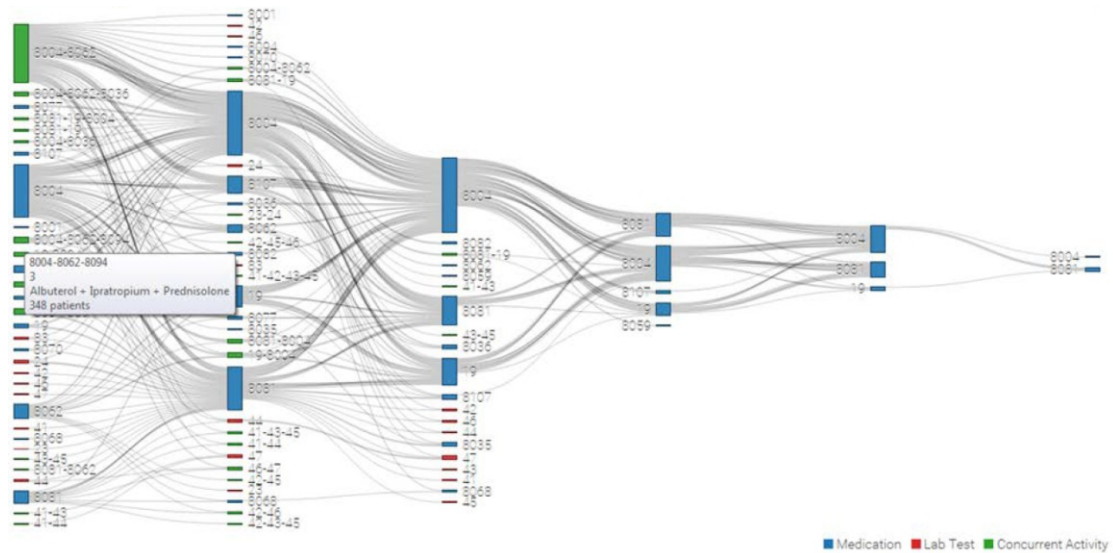


Figure 3.4: Web-based visualization [BPG<sup>+</sup>15]: Sankey diagram of subprocesses that visualize the flows between executed care activities.

The next visualization is a decision tree used for the visualization of the “Conformance”. The decision tree is based on the guidelines and visualizes all possible paths of the guideline as can be seen in Figure 3.5. Furthermore, it includes for each type of care process the number of patients that received this care process. The last part of this visualization is the conformance score which is shown at the end of each path indicating how strictly the guidelines were followed.

This approach was preliminarily evaluated by clinical collaborators who responded very well to the approach. Nevertheless, they had some improvement suggestions, for example, to add a visual query-based system. Furthermore, Basole et al. [BPG<sup>+</sup>15] state that they plan further user studies with clinicians and care quality managers.

### 3.2.2 Approaches visualizing cohort patient data

**Lifelines2** [WPQ<sup>+</sup>08; WPS<sup>+</sup>09] is the first approach in the category of approaches that deal with patient data visualization. The goal of this approach is to visualize event sequences to detect precursor, co-occurring as well as aftereffect events. The event sequences are visualized for multiple patients which are aligned on a common timeline



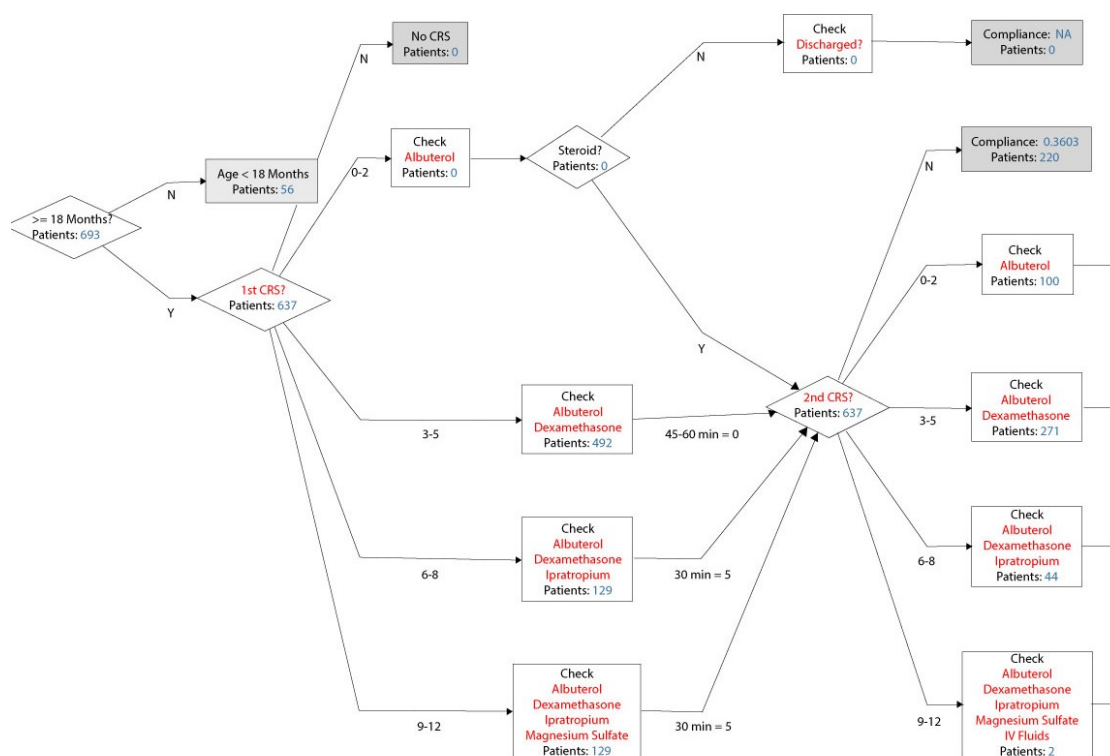


Figure 3.5: Web-based visualization [BPG<sup>+</sup>15]: Conformance flowchart visualizing all possible paths of the guideline. (Figure modified)

and are stacked above each other as can be seen in Figure 3.6. For the encoding of events, triangles are used that are differently colored depending on the event type. Interaction techniques, like being able to align, rank, or filter patients by a certain event are provided as well. The last part of the user interface is the temporal summary on the bottom.

“A temporal summary is a stacked bar chart that aggregates a variety of metrics (event counts, record counts, etc.) about a group of records over a time frame of varying granularities.” [WPS<sup>+</sup>09, p. 1051]

For the evaluation of their approach Wang et al. [WPQ<sup>+</sup>08] conducted two user studies where one of those was conducted with 20 non-medical experts and the other one with four medical experts. The intent behind the first user study was to evaluate the alignment functionality and its effectiveness and it turned out that the alignment was well accepted and seemed to be helpful. The intent behind the second user study was to receive feedback as well as improvement suggestions from medical experts. According to Wang et al. [WPQ<sup>+</sup>08] the medical experts quickly got used to the interface and stated to like the color-coding. Furthermore, the medical experts also stated that the visual representation of the clinical data could be improved a bit.

### 3. STATE OF THE ART

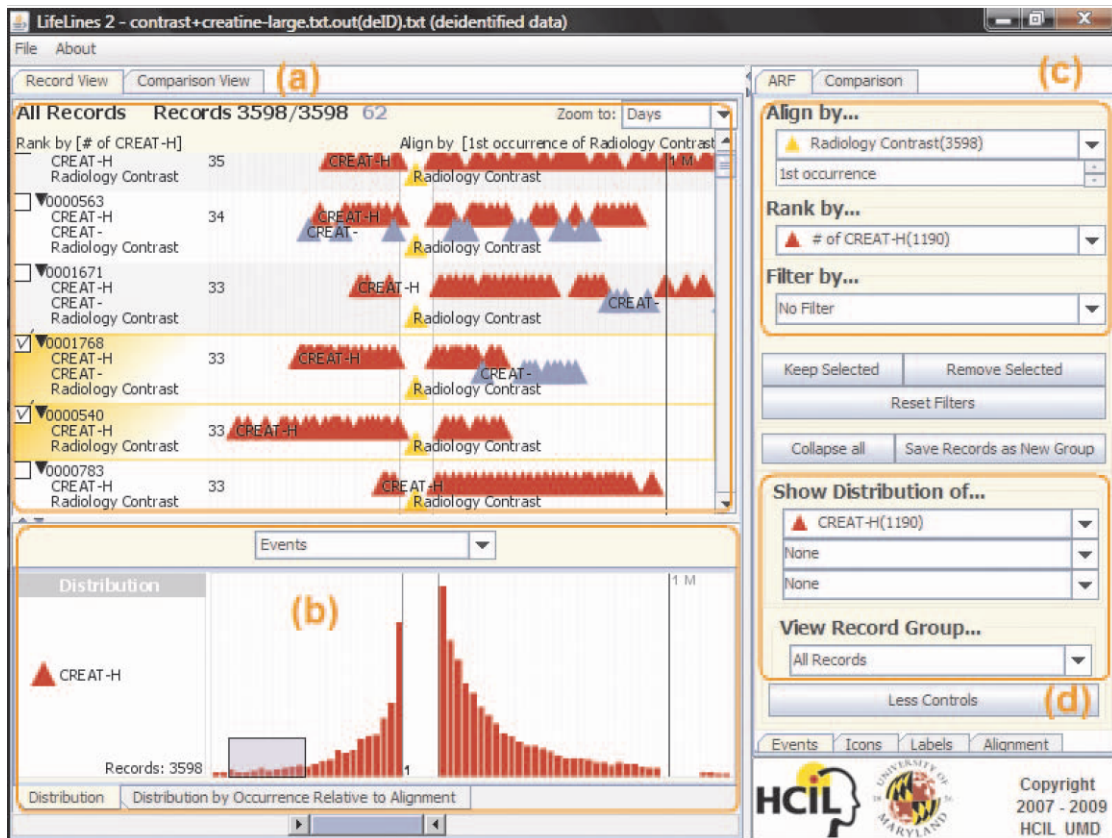


Figure 3.6: Lifelines2 [WPS<sup>+</sup>09]: An approach that visualizes event sequences of multiple patients. (a) shows the visualization of event sequences per patient and (b) a temporal summary. (c) are controls for alignment, ranking, and filtering, and (d) controls for the temporal summary and navigating groups.

The next approach in this category is called **LifeFlow** [WGP<sup>+</sup>11]. As can be seen in Figure 3.7 LifeFlow visualizes event sequences. Analysis of event sequences is for example important for medical researchers who want to study transfer patterns within a hospital. In Figure 3.7 it can also be seen that LifeFlow can be used in combination with Lifelines2 but a standalone version of LifeFlow is available as well.

The event sequences in LifeFlow are represented using a tree-based structure which is based on the prefixes of the event sequences meaning that sequences starting with the same event type are grouped and if the prefix changes the sequence is split. Each event is encoded by using a bar that is color-coded depending on the event type. The bar height encodes how many patients belong to the node. Furthermore, the distance between bars represents the mean time between events. The last feature is the same alignment feature as Lifelines2 uses which makes it possible to align the sequences by certain events.

Wongsuphasawat et al. [WGP<sup>+</sup>11] conducted two case studies as well as one user study.



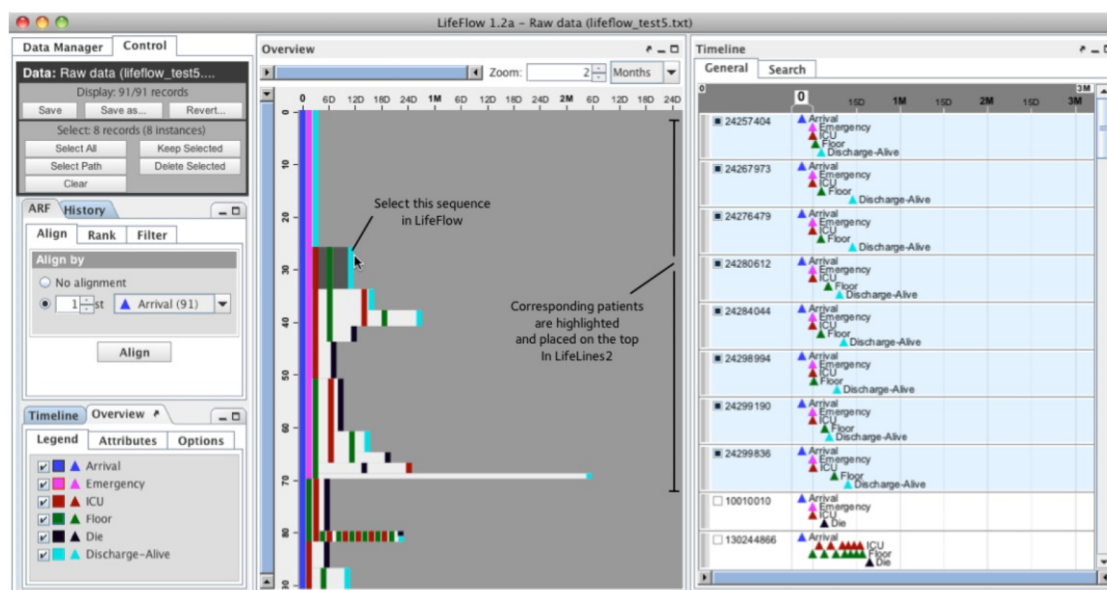


Figure 3.7: LifeFlow [WGP<sup>+</sup>11]: Visualization of event sequences. On the left, controls for aligning, filtering, and ranking are present as well as a legend for selecting certain event types. In the center, a tree-based visualization represents the event sequences and on the right, LifeLines2 is used for the visualization of event sequences of single patients.

For the first case study, they collaborated with a physician whose feedback was very positive. For the second case study, a different, non-medical dataset dealing with traffic incident data was used. The participants were domain experts who “stated that they wished LifeFlow was available earlier on when they started their own analysis” [WGP<sup>+</sup>11, p. 1754].

For the user study, ten graduate students were recruited to evaluate LifeFlow. Again very positive results were achieved since the participants got used to LifeFlow very quickly and had no major problems solving the tasks.

The **timeline-based visualization** from Dabek et al. [DJC17] is the last approach in this category. The goal of their system is to give an overview of patient histories. As can be seen in Figure 3.8 their approach is based on a vertical timeline. The timeline includes summary nodes which are added to the timeline for each date where an entry exists in the EHR of the patient. Furthermore, the summary nodes are enhanced by attribute nodes which represent the different actions that were performed on this day, like lab test results, medications, etc. To get more details about the record from a certain day they can be accessed by clicking on a summary node. For comparing multiple patients with each other a summary view is included where the timelines of multiple patients are stacked above each other.

The approach was not evaluated but it is stated that they are planning to conduct a case study with clinicians [DJC17].



Figure 3.8: Timeline-based visualization [DJC17]: Timeline-based visualization of a patient history.

### 3.3 Discussion

To summarize the results of our literature research we discuss in this section all the approaches we found and compare them with each other. Furthermore, their advantages and disadvantages are presented as well as their shortcomings.

We start with **CareVis** which is the first approach in the KAVA-related approaches category. The first positive aspect of CareVis is that a user-centered design approach was used. They collaborated with medical experts who supported them with domain knowledge which ensures that the approach covers all important aspects that physicians need for their work. Another positive aspect of CareVis is that during the evaluations its graphical user interface proved to be intuitive and was easily understandable for the physicians. Especially the QuickView panel was very well received since it provides a summary of the most important parameters. Furthermore, it was stated that with the PlanningLines approach temporal uncertainty can be visualized which is also useful.

However, the major drawback of CareVis is that only patient data from single patients is visualized which makes a comparison between patients difficult. Another drawback of CareVis is that highlighting how the executed actions have affected the patient parameters is only very sparse.

The next approach, **CareCruiser**, which is based on CareVis, solves these two problems. First, multiple patient records are stacked above each other which makes a comparison possible. Regarding the highlighting of the effects the actions have on the parameters CareCruiser added color-coding to make it more present and easy to detect. Another difference is that CareCruiser uses diamonds instead of PlanningLines for the visualization of the executed actions which is a very simple but intuitive approach. One advantage of the diamonds is that with the peak of the diamond it is possible to analyze the exact time when the action was executed. Furthermore, by using different colors for

the different action types they can be easily distinguished. To sum it up, CareCruiser is a well-developed approach that solves many of the drawbacks of CareVis. We could only identify one drawback which is the visualization of large patient cohorts. With CareCruiser it is possible to compare single patients with each other but if the number of patients gets larger it gets even harder due to the limited screen size.

Therefore, **Gnaeus** improved the patient data visualization view to be able to visualize data of large patient cohorts by aggregating the patient data. This enables getting an overview of a whole cohort very easily. Furthermore, Gnaeus includes multiple interaction techniques, like a grouping lense or a fish-eye distortion, to analyze the data of single patients as well.

When comparing all these three approaches it can be seen that they all use the same representation for the treatment plans which is a flowchart-like representation. An advantage of this technique is that flowcharts are well known to physicians which enables a fast learning curve [AM06]. Another advantage is that color-coding can be used in this technique which makes it possible to associate the actions in the treatment plan visualization with the ones that were executed.

The last approach in the category for KAVA-related approaches is the **web-based visualization** of Basole et al. [BPG<sup>+</sup>15]. Their approach enables the comparison of different patient cohorts as well as the analysis of guideline conformance. They use a Sankey diagram visualizing the mainly followed paths of executed actions which is a well-understood approach that makes the analysis of flows between certain subprocesses possible. One drawback of their approach is that the effects of the executed treatment plan actions cannot be seen directly.

**Lifelines2** is the first approach from the category of approaches that visualize only patient data and do not include any domain knowledge. The graphical user interface of Lifelines2 is based on a time-oriented visualization that is already well-known and, therefore, easy to understand without needing much learning effort. Another positive aspect of Lifelines2 is its align, rank, and filter functionality which was very well received during the user study. Furthermore, the temporal summary is very helpful for getting an overview of the distribution of certain event types. In total, Lifelines2 is a well-accomplished approach, only one aspect that could be improved is to add the possibility of being able to see the general flow of all event sequences.

This leads to the next approach, **LifeFlow**, which, in contrast to Lifelines2, visualizes the general flow of event sequences. Furthermore, LifeFlow can be combined with Lifelines2 which makes it possible to see, on the one hand, the general event sequences, and, on the other hand, the event sequences of single patients. Their evaluations have proved that LifeFlow is easy to use and needs only a low training effort.

The last approach is the **timeline based visualization** from Dabek et al. [DJC17]. With their approach, it is easily possible to get an overview of the medical history of a patient by visualizing many aspects of the EHR, like vitals, medications, etc. However, if one is interested in analyzing the data in more detail the effort increases due to the

### 3. STATE OF THE ART

complex structure and the large number of views which requires a lot of searching and switching between the views. Furthermore, there is the possibility of comparing the timelines of multiple patients with each other but since the timelines are only summaries a detailed comparison is more difficult.

In the following table (see Table 3.1) a tabular overview with all advantages and disadvantages of each presented approach is given.

Approach	Advantages	Disadvantages
<i>KAVA-related approaches</i>		
<b>CareVis</b> [AM06]	<ul style="list-style-type: none"> <li>easily understandable</li> <li>intuitive</li> </ul>	<ul style="list-style-type: none"> <li>no possibility of comparing patients</li> <li>only slight highlighting of the effects of the executed actions</li> </ul>
<b>CareCruiser</b> [GAK <sup>+</sup> 11]	<ul style="list-style-type: none"> <li>visualization of multiple patient records at once</li> <li>color-coding to highlight effects of executed actions</li> </ul>	<ul style="list-style-type: none"> <li>comparison and analysis of data from large cohorts difficult</li> </ul>
<b>Gnaeus</b> [FUA <sup>+</sup> 15]	<ul style="list-style-type: none"> <li>visualizes data of large patient cohorts</li> <li>different visualization types for the patient parameters</li> </ul>	
<b>Web-based visualization</b> [BPG <sup>+</sup> 15]	<ul style="list-style-type: none"> <li>comparing of different cohorts possible</li> </ul>	<ul style="list-style-type: none"> <li>effects of executed actions onto patient parameters not shown</li> </ul>
<i>Approaches visualizing cohort patient data</i>		
<b>Lifelines2</b> [WPQ <sup>+</sup> 08]	<ul style="list-style-type: none"> <li>graphical user interface is easily understandable</li> <li>well working align, filter, and rank functionality</li> </ul>	<ul style="list-style-type: none"> <li>visualization of general event sequence flow missing</li> </ul>
<b>LifeFlow</b> [WGP <sup>+</sup> 11]	<ul style="list-style-type: none"> <li>intuitive</li> <li>visualizes general event sequences flow</li> <li>combination with Lifelines2 possible</li> </ul>	
<b>Timeline-based visualization</b> [DJC17]	<ul style="list-style-type: none"> <li>good visual summary of a patients' EHR</li> <li>includes a variety of aspects included in the EHR</li> </ul>	<ul style="list-style-type: none"> <li>a lot of time and effort is needed to analyze the patient history in detail</li> <li>detailed comparison between patients difficult</li> </ul>

Table 3.1: Tabular summary of all advantages and disadvantages of each presented approach.

# Data Exploration using Tableau

After finishing the literature research the second part of the methodology “Data Exploration using Tableau” was conducted. In this chapter, the process, as well as the results and findings of this step, are described but before we start with those aspects a short introduction to the software Tableau<sup>1</sup> will be given.

## 4.1 Tableau

Tableau is a data visualization as well as Visual Analytics software that enables users to create meaningful visualizations fast and easily. Tableau is extremely powerful and flexible. Simple visualizations, like line charts or bar charts can be created with only a few clicks but even more complex visualizations, like parallel coordinate charts or radar charts are possible.

Furthermore, Tableau has a very convenient feature called “Show Me” that suggests visualizations that could fit the selected data fields. This feature is very helpful because it creates initial visualizations very fast which can then be the base for further adjustments or analysis.

Another feature we used during the exploration process are the so-called “Dashboards” in Tableau. With dashboards, multiple visualizations can be combined into one single visualization. The visualizations can then be aligned and arranged freely. Furthermore, filters from single visualizations are also included in the dashboard and can be applied to one or all visualizations. Another helpful feature of Tableau dashboards is that all the visualizations are connected and linking is, therefore, very easy. If a bar from a bar chart is selected the data that is connected to this bar gets highlighted in all the other visualizations where the data is included too.

---

<sup>1</sup><https://www.tableau.com/> (accessed 29/04/2021)

Therefore, Tableau dashboards are very helpful to create a whole interactive visualization including features, like filtering or linking. These are good to get an overview of a certain dataset easily which is why they were used in our data exploration phase.

### 4.2 Process

#### Data Cleaning and Preparation

In the beginning, the data were cleaned and prepared before used in Tableau. This preparation process included removing unnecessary data and also bringing the data into a format that works well with Tableau. For example, Tableau needs data that has just one label for each column. This was not the case for all files of our dataset. The file including the answers from the questionnaire for each patient and visit had multiple column labels and, therefore, we needed to remove some labels to obtain a file that fits the requirements of Tableau.

Furthermore, some information that was removed from the dataset is for example the date when the patient filled out the informed consent or many dates, like the birth date were saved in three different formats and we decided to keep only one format. To anonymize the data, all information that could connect the dataset to the specific study was removed.

The last step was to check if all files have a column containing the unique identifier of each patient and if the column has the same label in each file. The unique identifier is important for Tableau because when multiple files are imported Tableau wants to know based on which column the files should be merged.

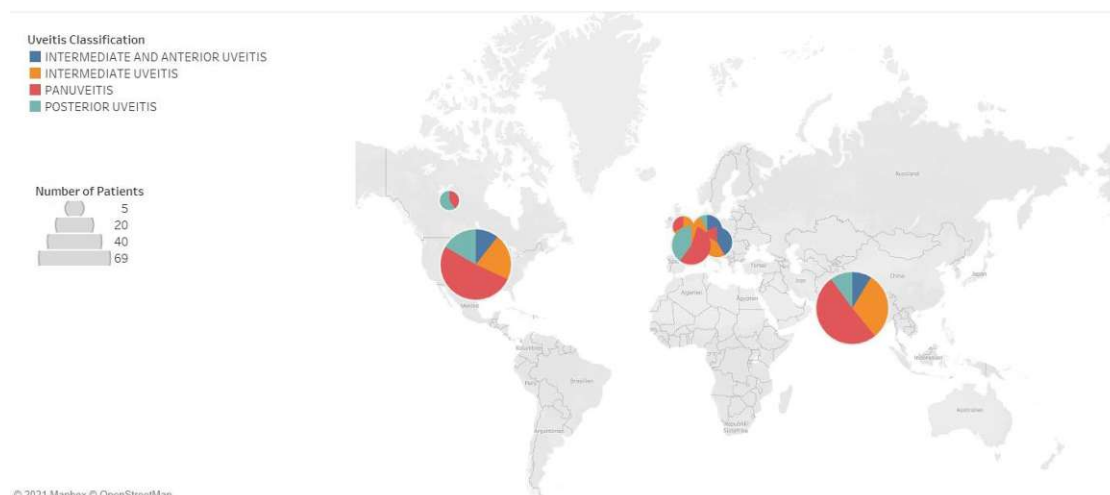


Figure 4.1: A world map visualizing the distribution of the various Uveitis classifications among the different countries.

After finishing the data cleaning and preparation process the data were imported into a Tableau project and merged based on the unique identifier and then first visualizations were created which are explained in the next section.

### Visualization of Demographic Data

The first visualizations created were rather simple and had the aim to understand the general patient data better and to see how the patients are distributed among the different demographic categories, like age, gender, country, etc.

The first category of demographic data that was visualized is the information about the country/region the patients were recruited from. For such information, a world map visualization suits very well and was also suggested by the “Show me” tool from Tableau. Therefore, a world map that shows how the different types of Uveitis are distributed among the different countries was created. For each country that is included in the dataset, a pie chart was added to the map. The size of the pie chart indicates the number of patients and the colors represent the different types of Uveitis. The result can be seen in Figure 4.1.

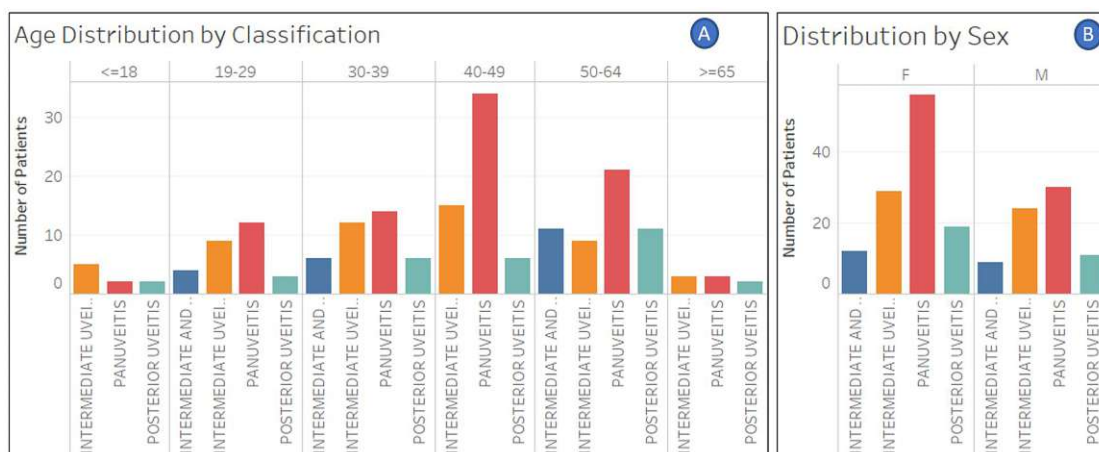


Figure 4.2: Bar charts visualizing distributions by Uveitis classification: (A) shows the distribution among the age categories and (B) shows the distribution among female and male patients.

To show the distribution of the patients among the categories, like gender, treatment group, age category, etc. bar charts were used. The shape of the bar chart is the same for all categories and the categories are easily exchangeable in Tableau. The bar chart shows the distribution of the category by the different Uveitis classifications. Examples of those bar charts can be seen in Figure 4.2 where (A) visualizes the distribution per age category and (B) the distribution per gender. As can be seen in the figure the bar charts show how many of the patients belong to which category and from each category how many patients belong to which Uveitis classification. This makes it possible to see,



on the one hand, which category the most patients belong to and, on the other hand, how the different Uveitis types are distributed between for example males and females or the age groups.

In this step of the exploration part, other types of visualizations were tried out as well, for example, treemaps or so-called “packed bubbles”.

Those visualizations can be used to group the patients to maybe find patterns or groups of patients with similar characteristics that can be a base for further investigations at later stages. An example of the “packed bubbles” visualization is presented in Figure 4.3.

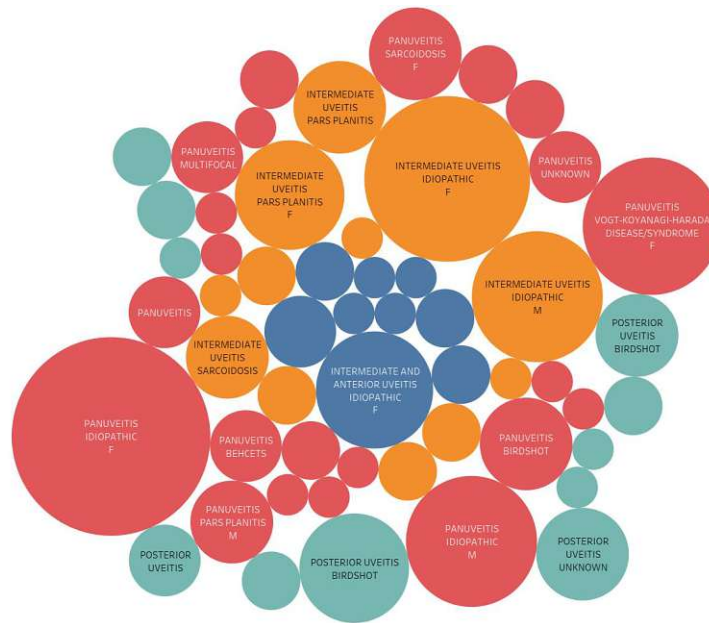


Figure 4.3: “Packed bubbles” visualizing groups of patients based on the Uveitis classification, etiology, and gender.

After finishing the visualizations for the more general data, the questionnaire data were explored which is explained in the next section.

#### Visualization of Raw Questionnaire Data

The second important part of the dataset is the data regarding the two questionnaires that give information about the quality of life of the patients. As already mentioned the patients answered those questionnaires multiple times to see if and how the answers changed over time.

The amount of questionnaire data that was produced is much denser than the personal data from the patients. Therefore, visualizations, like bar charts that we used before are not suitable because of the resulting visual clutter and the limited screen size. This means more compact visualization types are needed and for this task, the “Show me” tool



was again helpful. Tableau suggested for example heat maps where the x-axis represents the patients that have a column for each visit and the y-axis represents the questions from the questionnaire. An example of a resulting heat map for the SF36 questionnaire can be seen in Figure 4.4.



Figure 4.4: Heat map for the SF36 questionnaire answers of all patients per visit.

Those heat maps were a good starting point because it was possible to see all the data at once and with the use of filters the heat map becomes easier to read. However, this kind of heat map was not the best solution because it presents all the information at once which makes it hard to read and it does not help to find any patterns or insights. Such heat maps are more helpful for example if you already know what kind of questions or patients you are interested in and then it can be used to look up how the patients answered certain questions but for exploration tasks, it is more difficult because of the huge amount of data that is visualized at once.

Therefore, the next step was to find other visualizations for the questionnaire data. To find proper visualizations Miksch and Aigner [MA14] recommend to first define the data, user, and tasks for the visualization. Their data-user-task triangle will be used at a later point in our thesis to a greater extent as well. However, in the data exploration phase, we defined general tasks to find proper visualizations for the questionnaire data. The main task we defined is that it should be possible to find out if the patients' answer for each question improved or worsened across all visits because, in the end, it is important to see if the treatment helped to improve the quality of life or not.

Based on this task we tried to find proper visualizations and thought about encodings that could fit and the result were two different visualizations. The first one is a sparkline chart which can be seen in (A) in Figure 4.5. Sparklines, also known as word graphics, are

graphics in the size of words and usually in the form of line charts but without displaying the axis [GWFI14].

For the questionnaire data, a sparkline for each question and patient was created that displays the temporal evolution of the answers per patient. The advantage of this encoding is that it is space-efficient and a trend if the patients' answer improved or not, can be seen very easily. A minor drawback is that the exact values of the answers are not displayed which makes it difficult to distinguish between straight lines that indicate that the answer did not change. Furthermore, if the improvement or worsening was rather low from visit to visit it is difficult to analyze as well. One possible solution to better distinguish between the lines would be to use color-coding as in the heat map.

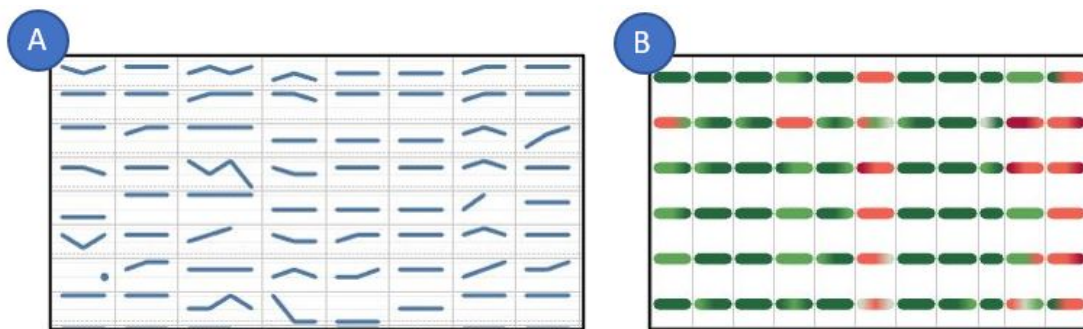


Figure 4.5: Questionnaire data encodings: (A) Sparkline chart, (B) Color-coded lines.

Based on these concerns the second idea evolved which can be seen in (B) in Figure 4.5. This encoding is similar to the sparkline approach because it is also split by patient and question but instead of drawing a line chart, a straight line is drawn for each question and each patient. This line has a thicker stroke and is divided into multiple sections where each section represents one visit and the section is colored using the same color-code as in the heat map in Figure 4.4. Dark red means the answer was 0 and dark green means 100. This visualization has the advantage that it needs even less space than the sparklines and the problem of not being able to distinguish between lines that did not change over time is solved by using color-coding.

However, those two visualizations just used different types of encoding for the data but the amount stayed the same and although those two solutions are rather space-saving it was still too much information that was displayed at once.

Therefore, another solution with aggregated data was developed. Based on the previously defined task the data were aggregated in a way that for each patient and question a value that indicates whether the answer of the patient improved or worsened throughout the visits was calculated. This aggregated value is calculated by using the difference of the answer value of the last visit and the first visit and map this value to a range from -2 to +2. This resulted in a reduced dataset where each patient has only one value per question which was then again used as an input for a heat map visualization as can be

seen in Figure 4.6. The color-coding as displayed in the legend ranges from dark red to dark blue where dark red means that there was a significant worsening from the first visit to the last visit and dark blue indicates a significant improvement.

The combination of the aggregated dataset and the heat map visualization gives a good overview of how the patients' answers have changed over time and outliers or special cases can be found more easily. However, the heat map itself without providing any context information would not be sufficient and interactivity is needed. One interactive part are the filters on the right that could help to reveal any patterns for example by filtering only for patients that got the placebo and to analyze if there were more patients with significant worsening than in the other groups. Furthermore, in Tableau, it is possible to add tooltips to the visualization to provide more details about the underlying data. Therefore, we defined what kind of information regarding the patient should be displayed on hovering over a data point in the heat map. The result can be seen in Figure 4.6.



Figure 4.6: Heat map visualization of the aggregated data for both questionnaires.

However, these visualizations above were all based on the raw questionnaire data and the aggregated data but there is also another aspect of the data which was not observed yet. As already discussed the two questionnaires both provide multiple subscores and the next step was the exploration of the subscores which is discussed in the next section.

### Visualization of Questionnaire Subscores

For the subscore visualization, we again defined certain tasks as a starting point for the development of the visualizations. The first task was that the visualization should help to explore and compare the subscores between the different groups of patients. For example,

#### 4. DATA EXPLORATION USING TABLEAU

to be able to analyze if there are any differences between males or females or between the different age groups. Based on the data Tableau's "Show me" feature suggested bar charts as can be seen for example in Figure 4.7 where the subscores from males and females are compared. Multiple bar charts were created for the different categories, like age groups, Uveitis classification, etc. to see if there are any patterns or any differences between certain groups of patients.

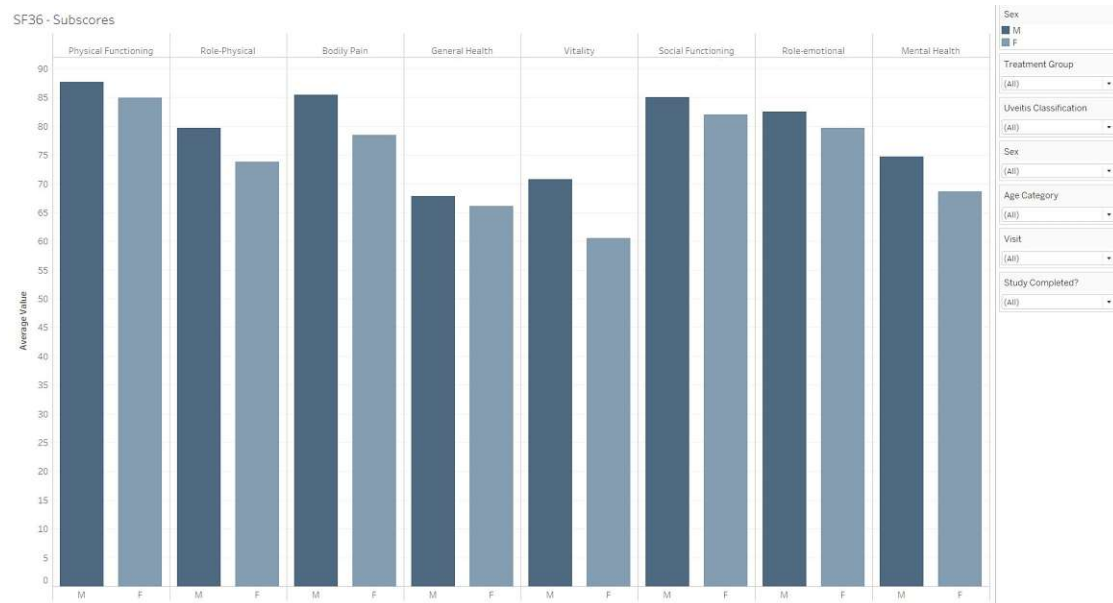


Figure 4.7: Bar chart visualizing the subscores from the SF36 questionnaire for both male and female patients.

After this part of the exploration was finished the second task was defined which deals more with the temporal progress of the subscores. This means that the visualization should help to get to know how the subscores have changed from visit to visit to maybe find certain subscores that improved very well or worsened. Such findings could then be used for further investigations and exploration.

However, for this task, we did not use the "Show me" feature from Tableau. During our research, we found multiple papers that dealt with Uveitis and quality of life using the same questionnaires and they tried to visualize these data as well. One example is the paper from Fabiani et al. [FVO<sup>+</sup>17]. Their study's objective was "to investigate the impact of non-infectious uveitis (NIU) on quality of life (QoL)" [FVO<sup>+</sup>17, p. 478]. They had 80 NIU-patients and 23 healthy controls and used the SF36 questionnaire to investigate if there are any differences between the NIU-patients and the healthy controls and if correlations between the subscores and the demographic data can be found. To analyze the data they gathered, they used visualizations, like radar charts and bar charts [FVO<sup>+</sup>17].



Therefore, we decided to use the same types of visualizations for our dataset because they proved to be effective for the task. Especially the radar chart seems to be an effective visualization technique for data from clinical studies as Saary [Saa08] concludes it in her paper as well. Radar charts also called radial plots are a circular graphing method for multivariate data. Furthermore, Saary [Saa08] proposes some example use cases, like the visualization of temporal changes in multiple variables or to analyze if there are differences between patient groups concerning multiple measures which are both use cases that fit perfectly to our dataset and our predefined tasks.

Another reason why we decided to use the same visualizations is that it enables us to compare the results from Fabiani et al. [FVO<sup>+</sup>17] with ours and maybe find some similarities.

Nevertheless, it needs to be considered that their study setup was a bit different than ours which makes the comparison not possible in some cases. For example, they compare NIU-patients with healthy controls whereas in our dataset healthy controls are missing and their participants answered the questionnaire only once while we have data from multiple visits.

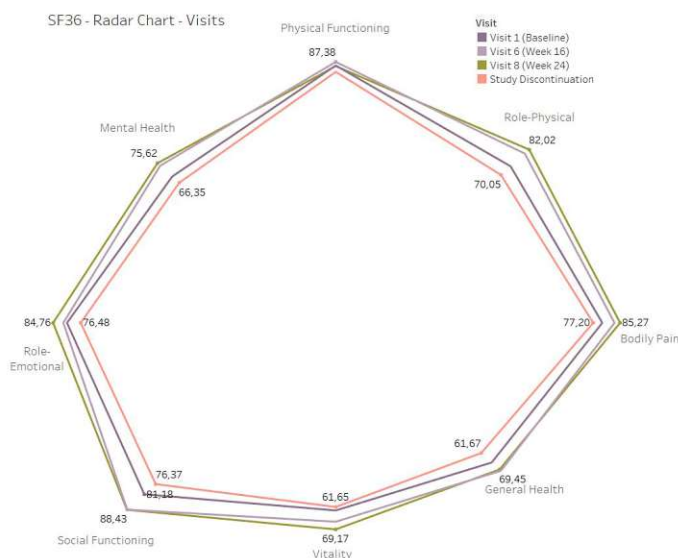


Figure 4.8: Radar chart visualizing the subscores from the SF36 questionnaire per visit.

As we already created bar charts before and also compared males and females as they did in their paper the next step was the creation of the radar chart. In the paper, the radar chart was created for the comparison of patients with active Uveitis and patients with inactive Uveitis. As already mentioned our dataset is a bit different and, therefore, we decided to use the radar chart to compare the subscores from the different visits. The radar chart for the SF36 questionnaire subscores can be seen in Figure 4.8 where each line represents one type of visit and each corner corresponds to one subscore. Usually,

radar charts are drawn based on a spider net but with Tableau, it was easier to create this kind of radar chart and as the focus of this whole process is more on exploring the data instead of creating sophisticated charts we stayed with this type of radar chart.

The advantage of a radar chart is that it is possible to analyze if for example, the subscores improved over time or if certain subscores worsened. Furthermore, in combination with filters, certain groups of patients can be compared, for example, to analyze how the subscores from the placebo group changed from visit to visit in contrast to one of the other treatment groups or to compare the different age groups, etc.

Finally, after creating and completing all charts the next step was to combine certain charts into dashboards where each dashboard has a different goal or information that it should convey and the final results are described and presented in the next section.

### 4.3 Final Dashboards

The first dashboard that was created which can be seen in Figure 4.9 is based on the visualization of the demographic data discussed in Section 4.2. The goal of this dashboard is to give an overview of the patients that participated in the study, which characteristics they have, and to see how the patients are distributed among the different categories. Furthermore, it could help to identify certain groups of patients with similar characteristics. Such groups can then be used as an input for further exploration tasks in other dashboards by using filters and being able to compare certain groups in order to get more insights into the data.

The dashboard consists of four visualizations, multiple filters, and a small table. (A) shows the world map that was already discussed in Section 4.2 whose goal is to visualize the distribution of the different Uveitis classifications among the countries present in the dataset. The bar charts (B)-(D) show how the different Uveitis classifications are distributed regarding gender, age categories, and etiology.

Furthermore, as already mentioned dashboards in Tableau are interactive. In Figure 4.9 (E) shows some filters, like a filter for patients that completed the study or a filter for patients from certain treatment groups, etc. Those filters affect all visualizations on the left. Furthermore, linking is possible, for example, the Panuveitis bar can be selected in one of the bar charts and then all other charts highlight the Panuveitis part as well. These features can help to explore the data even more and in greater detail.

With Tableau, it is also possible to aggregate the data and show it in a tabular form as can be seen in (F) in Figure 4.9. This table gives information about how many patients belong to which treatment group but this can be changed as well to for example another characteristic, like the race or age group if needed.

For the visualization of the questionnaire data, another dashboard was created which can be seen in Figure 4.10. This dashboard consists of two heat maps, a filter, and a legend pane. The intent behind this dashboard is to visualize the raw questionnaire data

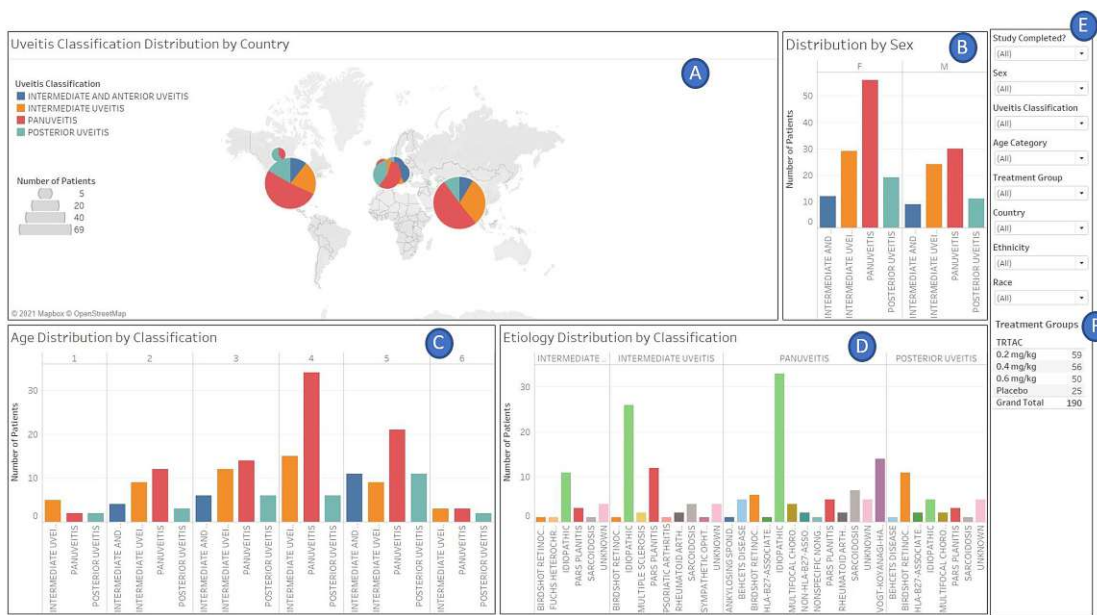


Figure 4.9: Dashboard visualizing general patient data: (A) world map visualizing the Uveitis classification per country, (B)-(D) bar charts visualizing the Uveitis distribution per gender, age category, and etiology, (E) filter pane, (F) table displaying patient distribution per treatment group.

to achieve an overview of the questionnaire data as a whole and in combination with the aggregated data, it is possible to find certain patients that for example improved from visit to visit and those can then be analyzed in more detail.

In the first heat map (A) each row represents a patient and each column represents a question from one of the two questionnaires (SF36 and VF25). As already explained in Section 4.2 this heat map visualizes how the answers of each patient have changed over the three visits. The change has 5 steps where the “worst” one is colored in dark red which means that there was a negative change from the first to the last visit and the “best” one is colored in dark blue which means that there was a positive change.

The second heat map (B) shows the detailed answers from each patient for each question and each visit. In this heat map, the color scale ranges from dark red to dark green where dark red means the patient answered the question with 0 and dark green means 100.

As with the first dashboard, this dashboard is interactive as well. You can for example hover over a patient in the first heat map (A) and more details about the patient are displayed in a tooltip. Furthermore, you can select a patient in this heat map and then the second heat map (B) displays only the answers from the selected patient. The last part of the dashboard is the filter and legend pane (C) which includes again multiple filters that affect all visualizations from the dashboard and the color legends from the two heat maps.

## 4. DATA EXPLORATION USING TABLEAU

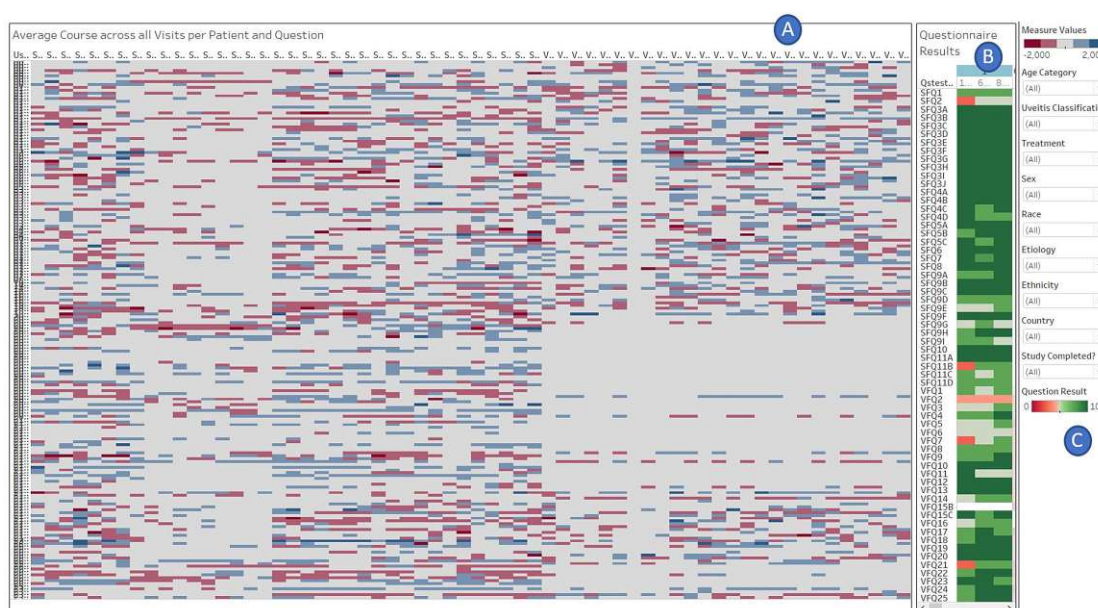


Figure 4.10: Dashboard visualizing questionnaire results: (A) heat map visualizing the change from the first visit to the last visit, (B) heat map visualizing the answers of all visits, (C) filter and legend pane.

The two final dashboards deal with the questionnaire subscores and are based on the visualizations from Section 4.2. Two separate dashboards were created for each of the questionnaires to give an overview of the subscore data.

The dashboard for the SF36 questionnaire can be seen in Figure 4.11. The first two bar charts (A and B) visualize the summary scores Mental Component Summary (MCS) and Physical Component Summary (PCS) split by gender (A) and by Uveitis classification (B). The next bar chart (C) is the one already discussed in Section 4.2 where the subscores from females and males are compared.

Furthermore, the radar chart (D) visualizes the subscores split by visit to analyze if there were any remarkable changes between the visits. The last chart is again a bar chart (E) visualizing the subscores split by Uveitis classification to check if there are any differences between the Uveitis types regarding the scores. For example, it could be that a certain Uveitis affects the mentally related scores more than the physical ones or the other way round and this could be analyzed with this chart.

The last part of the dashboard is again a filter and legend pane (F). Interactivity and linking are available as well in this dashboard.

The dashboard for the VF25 questionnaire is very similar to the SF36 questionnaire dashboard and can be seen in Figure 4.12. The only difference is that the VF25 questionnaire does not have summary scores and, therefore, the two bar charts for the



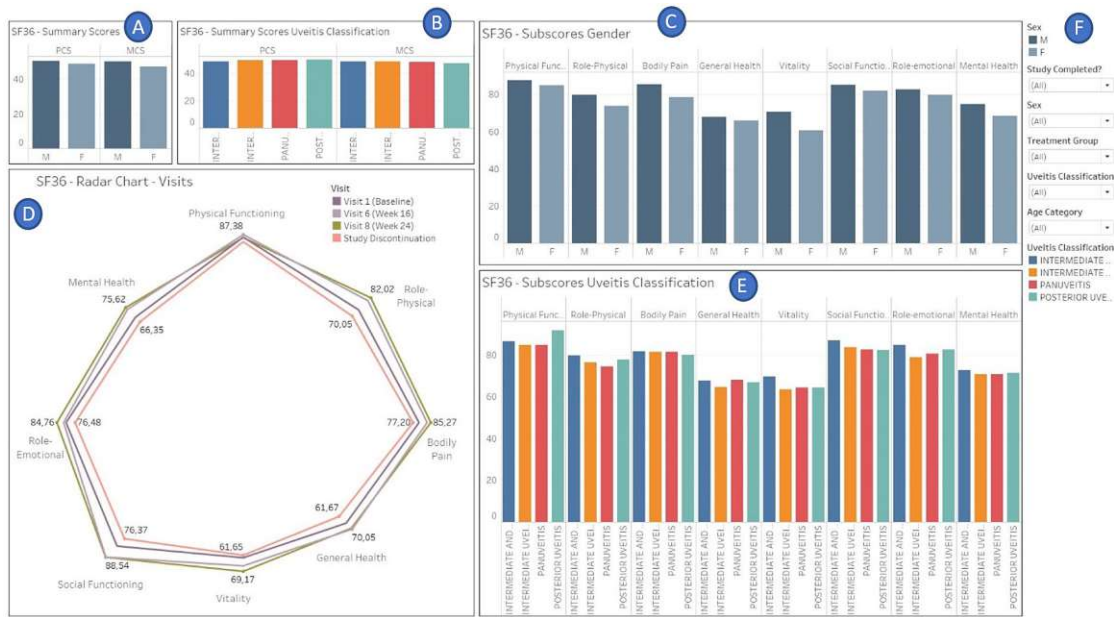


Figure 4.11: Dashboard visualizing subscore data from SF36 questionnaire: (A) bar chart visualizing summary scores per gender, (B) summary scores per Uveitis classification, (C) bar chart for subscores per gender, (D) radar chart visualizing subscores for each visit, (E) bar chart visualizing subscores per Uveitis classification, (F) filter and legend pane.

summary scores on top of the radar chart were removed. All the other charts, filters, and interactive features are the same as in the SF36 questionnaire dashboard.

## 4.4 Data Exploration Findings

As a last step of the exploration part, the different visualizations and dashboards were used to analyze the data and the results are presented in this section. The findings are divided into the same subsections as in Section 4.2.

### 4.4.1 Demographic Data

First, the demographic data were analyzed with the intent of finding relevant information about the dataset, for example, to detect groups of patients with similar characteristics that are more present than others. Those groups can then be used for further analysis to see if their questionnaire results are similar or if there are any differences, etc.

The first finding is that more female patients (116 patients) participated in the study than males (74 patients). From the six different age groups, the group 40-49 with 55 patients (29%) and 50-64 with 52 patients (27%) are the ones with the most patients. The least patients are younger than 19 (5%) or older than 64 (4%). Regarding the different

#### 4. DATA EXPLORATION USING TABLEAU

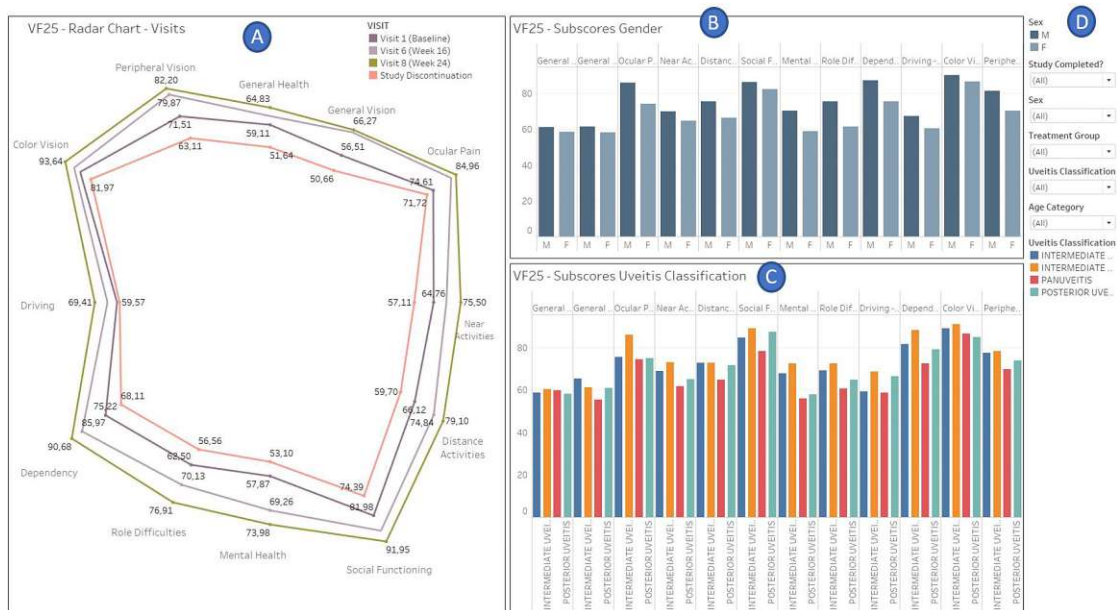


Figure 4.12: Dashboard visualizing subscore data from VF25 questionnaire: (A) Radar chart visualizing subscores per visit, (B) bar chart visualizing subscores per gender, (C) another bar chart that visualizes the subscores per Uveitis classification, (D) filter and legend pane.

countries, the most patients are from India (36%) and the USA (34%) and the least from Canada (3%) but concerning the region, the least patients are from Europe (26%).

The next important information is the Uveitis classification and etiology. Panuveitis is with 86 patients (45%) the most common for both males and females followed by Intermediate Uveitis with 53 patients (28%) and only 21 patients (11%) are affected by Intermediate and Anterior Uveitis. The most common etiology is “Idiopathic” except for Posterior Uveitis where “Birdshot Retinochoroidopathy” is the most common etiology. These findings are all based on the bar chart and world map visualizations from the dashboard shown in Figure 4.9.

Using the “packed bubbles” visualization from Figure 4.3 some groups of patients with similar characteristics were identified, for example, 23 female patients (12%) are affected by Panuveitis with “Idiopathic” as etiology and another group of 16 female patients (8%) with Intermediate Uveitis and also “Idiopathic” as etiology. Whereas 10 male patients (5%) are affected by Panuveitis with again “Idiopathic” as etiology and another 5% of male patients with Intermediate Uveitis and “Idiopathic” etiology.

The last aspect we investigated was the different treatment groups and we discovered that in the placebo group the age groups 1 ( $\leq 18$ ) and 6 ( $\geq 65$ ) are not represented whereas all other treatment groups include patients from all age groups.

All in all, the exploration of the demographic data was a good starting point to gain insights into the patients and their characteristics. The dashboard as well as the other visualizations that were created before simplified the exploration task even more and were a great help. Even though the single visualizations were not very complex or sophisticated the combination of all of them enhanced with filters and interactivity, like linking made them very effective for our task.

#### 4.4.2 Raw Questionnaire Data

As already discussed, the idea behind the heat map visualizations of the raw questionnaire data as well as the aggregated data as seen on the dashboard in Figure 4.10 was, on the one hand, to get an overview of the whole questionnaire data and, on the other hand, to maybe find questions or patients/groups of patients whose answers significantly worsened or improved to find patterns.

However, the first point that we discovered based on the visualizations is that from the 69 patients from India only 9 patients filled out the VF25 questionnaire and for all the others the data is either missing or they did not fill it out. Whereas the SF36 questionnaire was filled out by all patients except for one patient from the USA and a few patients did not answer the SF36 questionnaire at all visits.

In general, it can be said that based on the aggregated values most of the answers did not change significantly from the first visit to the last as the number of grey rectangles are prevalent in the heat map.

Focusing on the different treatment groups we analyzed that in each treatment group except for the placebo group more than 50% of the patients dropped out earlier from the study. From the placebo group, 12 out of 25 patients (48%) dropped out earlier and for the other three treatment groups, the drop-out rate is between 57% and 58%. Furthermore, the placebo group is the one with the least number of significant worsening (colored as dark red in the heat map) followed by the 0.6mg/kg group and the 0.4mg/kg group is the one with the most significant worsening. However, observing the significant improvements (colored as dark blue in the heat map) the 0.4mg/kg is the one with the most improvements followed by the 0.2mg/kg and 0.6mg/kg groups.

Something that can be observed as well is that if single patients significantly improved or worsened in a question then they often have other questions that improved or worsened significantly too. For example, in the placebo group only four significant worsening and nine significant improvements can be noted and from those one single patient has three significant improvements and two significant worsenings. However, this can be of course a coincidence but it is something that could be observed in more detail.

We also analyzed if certain questions stand out, for example by having multiple patients that significantly improved or worsened but we could not find any patterns. The question with the most significant improvements (three significant improvements) is VFQ3 with the question text “How much of the time do you worry about your eyesight?” but with a

count of three, it is still not much. Similarly, the question SFQ9I (“Did you feel tired?”) with a total of four significant worsenings is the one with the most worsenings but out of 190 patients, this only amounts to 2% and is, therefore, not significant.

Furthermore, in the 0.4mg/kg group we could observe two patients that stood out. The first one is a 50-years old female from the USA affected by Posterior Uveitis who had many good scores compared to the others from the group. However, she still dropped out after the sixth visit. Something interesting is also that she had multiple significant improvements but all were only noticed in the SF36 questionnaire. The second patient from the same treatment group had significant improvements in six questions but all these questions were from the VF25 questionnaire whereas the results from the SF36 questionnaire were not that good with only one slight improvement and more slight worsenings or constancies.

While these observations only regard two patients, making them not significant, it could still be worth conducting a more in-depth comparison between the two questionnaires.

Nevertheless, in the end, we could find some interesting aspects and the heat map visualization with the aggregated data was very effective for the exploration task. The heat map itself is in general a very dense visualization technique which makes it possible to visualize a greater amount of data at once and due to the aggregated data it was possible to find certain questions or patients that stood out by a significant improvement or significant worsening very fast. Furthermore, the filters proved to be very effective for comparing certain groups of patients. All in all, the dashboard was a great help for the exploration of the questionnaire data.

### 4.4.3 Questionnaire Subscores

Finally, the subscore data from the questionnaires were analyzed and for this part, two different approaches were used. The first one is that as already mentioned we compared our results and visualizations with the ones from the paper from Fabiani et al. [FVO<sup>+</sup>17] and the second one is the analysis of the visualizations where a comparison was not possible because of the different setups of our study and their study.

Therefore, we start with the comparison where we found some similarities as well as some differences. One similarity is that in both datasets the subscores “General Health” and “Vitality” are the ones with the lowest scores. “Physical Functioning” is in both datasets one of the highest scores whereas in our dataset “Bodily Pain” is rather high compared to their score.

Comparing the subscores from males and females in both datasets the subscores from the female patients are always lower than the ones from the male patients except for “General Health” in our dataset. Furthermore, the biggest difference between males and females is in the “Bodily Pain” subscore in our dataset and in their dataset, it is “Role-Physical” and “Bodily Pain”.

The last aspect that was compared are the summary scores PCS and MCS which are below 50 in both datasets but ours are closer to 50 than theirs. Furthermore, in our dataset PCS is higher than MCS whereas in their dataset it is the other way around. Comparing PCS and MCS between males and females again in both datasets the female scores are lower but the difference in both datasets is lower than 5.

Based on these findings we can say that there are some similarities but it needs to be considered that their setup was different than ours but it helped for the exploration and analysis of our data to see what visualizations they used and which aspects they observed.

However, as already mentioned, we did some other analyses as well especially regarding the temporal changes from visit to visit and the subscores from the VF25 questionnaire needed to be analyzed too.

We start with the two remaining charts from the SF36 questionnaire dashboard (see Figure 4.11) which are the radar chart, and the bar chart comparing the subscores for each Uveitis classification. In the bar chart, no huge differences between the different Uveitis types can be found. The highest subscore is “Physical Functioning” for Posterior Uveitis and the lowest one is “Vitality” for Intermediate Uveitis but all in all the differences are not huge.

However, the radar chart revealed some aspects for example that the “Study Discontinuation” visit has in each subscore the lowest score which makes sense because in most cases “Study Discontinuation” means that the patient needed to drop out because their health worsened. Furthermore, in all subscores, there was an improvement from the first visit to the last visit. This could indicate that the medication the patients received helped but this needs more detailed research. The differences from visits 6 and 8 were in many cases rather low but still improved except for “Physical Functioning” where visit 8 got a lower subscore than visit 1 but the difference is here low as well.

The last dashboard we analyzed is the dashboard for the VF25 questionnaire subscore data (see Figure 4.12). From the VF25 subscores, “Color Vision” is the one with the highest score followed by “Social Functioning” and the lowest is “General Vision” followed by “General Health”. Considering the gender it is here again as with the SF36 subscores that the female scores are always lower than the ones from the males. Furthermore, the score with the biggest difference between the male and female scores is “Role Difficulties”.

However, comparing the different Uveitis classifications two stand out which are Intermediate Uveitis which is the classification with the highest scores in 10 out of 12 scores and Panuveitis is the one with the lowest scores in 10 out of 12 scores.

As the last step, the radar chart was analyzed for the VF25 questionnaire as well. Again “Study Discontinuation” is the one with the lowest values in all subscores. Furthermore, there is always an improvement for each subscore from visit to visit.

Analyzing the placebo group in more detail we discovered that all patients that dropped out of the study already dropped out after the first visit. In all the other treatment

groups we could not find any significant aspects. In some groups, there was not a steady improvement in all subscores from visit to visit but most of the time only two subscores were affected.

Again we could gain new insights into the data with the help of the two dashboards. Especially the radar chart was effective for the exploration of the temporal progress from visit to visit because it can be seen at a glance if there was steady progress through the visits or not by searching for crossings between the lines. Moreover, big gaps between lines or points that are rather low or high can be a base for further analysis as well.

To sum it up, these were the main findings from the second part of our methodology “Data Exploration using Tableau”. In general, this step of the thesis was very helpful to understand the data better. The software Tableau was the ideal tool for this task since it is easy to use, proposes possible visualizations, and includes interactivity. We gained many insights into the data and the findings and experiences with the dataset as well as the different types of visualizations provide a good base for our next step which is the “Design and prototypical implementation of a new system”. The process as well as the results of this step are covered in the next chapter.

# Design Process

The next step after finishing the data exploration was the design of an own approach and in the following chapter, the process of this phase is presented. The base for this approach are the findings and insights from the previous phases. As already mentioned in Section 2 Methodology, the design process consists of multiple steps and is based on a user-centered design process in combination with problem-driven visualization research.

Therefore, the first step was the problem characterization which is discussed in the next section. After finishing the problem characterization, mockups were created that were constantly adapted and improved based on the feedback from visualization experts as well as the domain expert (see Section 5.2). In the end, after multiple feedback iterations, a final design for the approach was developed which was then prototypically implemented.

## 5.1 Problem Characterization

The first step, the problem characterization is, on the one hand, based on the data we received, and, on the other hand, based on the requirements from the domain expert. Initially, the data-user-task triangle from Miksch and Aigner [MA14] was used which “is intended to ease the design process and covers data, users, and their tasks” [MA14, p. 1]. Furthermore, the definition of those three components helps to find out how the data should be visually represented and which interaction techniques are suitable [MA14].

### 5.1.1 Data-User-Task Triangle

In this section, the three components from the design triangle from Miksch and Aigner [MA14] are described.



### Data

As already mentioned in the beginning, the basis for this thesis and the resulting prototype is a dataset dealing with the relatively uncommon eye disease called “Uveitis”.

The dataset itself resulted from a study that assessed the efficacy of a certain medication as treatment for Uveitis patients. The study itself lasted for six months and 190 patients with one of the following types of Uveitis participated:

- Intermediate Uveitis
- Intermediate and Anterior Uveitis
- Posterior Uveitis
- Panuveitis

Furthermore, the patients were split into four different treatment groups. Three of them got different doses of the medication and the last group received a placebo. The patients had their health state checked by a physician multiple times throughout the six months. At three of those visits, the patients were asked to fill out two different questionnaires as well. The questionnaires aimed to assess the current quality of life of the patient. The results of those questionnaires are also part of the dataset.

Therefore, the dataset itself consists of three parts. The first part contains patient-related data, like age, gender, ethnicity, race, height, weight, the country they are from and, data related to their disease. It includes the Uveitis classification, the etiology, and the affected eye.

Furthermore, it includes information about the visual acuity of the patients. The visual acuity was measured using LogMAR charts. First, the patients were asked to read the chart from a 4-meter distance and the number of letters the patients were able to read was documented. If the patients were not able to read a certain number of letters they needed to perform another test at a 1-meter distance. If the patient still was not able to read a certain number of letters a last test was performed. It was first tested if the patient could recognize finger counting, if not, hand motion recognition was tested. If this was also not possible the recognition of light perception was tested and if this was still not possible the patient was classified as blind. The results of all of these tests are part of the dataset as well as a score that summarizes all these tests and represents the visual acuity with a single value.

The second part of the dataset is the data from the questionnaires. For each patient and each visit, the answers for both questionnaires (if available) are documented. The last part is the subscore data from the questionnaires which is explained in detail in the next section. As already mentioned in Section 4 the questionnaire data was aggregated in a way that each patient got a value assigned for each question that indicates whether the



patients' answer has improved or worsened over the visits and this data is part of the dataset as well.

All in all, the dataset consists, on the one hand, of quantitative data that is time-oriented and linearly arranged, and, on the other hand, of qualitative data, like age, gender, race, etc.

### Questionnaires

As stated above the patients filled out two different questionnaires to assess the quality of life at three of their control visits. Those two questionnaires are the Visual Function Questionnaire (VF25) and the Short-Form 36 Questionnaire (SF36).

#### *Visual Function Questionnaire (VFQ-25)*

The VF25 is a questionnaire developed by the National Eye Institute (NEI) and is a shorter version of the 51-item VFQ containing only 25 questions. The VF questionnaire assesses how vision influences the quality of life including aspects, like emotional well-being, social functioning, or general health [Col02].

The 25 questions included one question for rating the general health and all the others represent 11 vision-related subscores, like a global vision rating, a mental health score or a score for color vision, etc. [Man].

For the calculation of these subscores, a two-step process needs to be conducted which is described in detail in the manual of the questionnaire [Man]. First, the answers from each question need to be re-coded based on predefined scoring rules and converted to a scale that ranges from 0 to 100 where 0 is the lowest score and 100 the highest. In the second step, the final calculation of the 12 subscores is done by averaging the scores from step 1 based on a list of questions that informs about which question contributes to which score. For example, for the calculation of the Ocular Pain score, the average of two questions (4 and 19) needs to be calculated.

#### *Short-Form 36 Questionnaire (SF36)*

The SF36 questionnaire consists of 36 questions and was initially developed for the Medical Outcomes Study (MOS) to examine the health status of the participating patients [Fra92]. The MOS was an observational study lasting for two years where patients with chronic conditions participated. The study had two goals where one of the two goals was the development of tools that enable the monitoring of patient outcomes as well as their determining factors. Therefore, one of the resulting tools was the SF36 questionnaire for the assessment of the quality of life [Wha13].

As with the VF25 questionnaire for the SF36 questionnaire, subscores exist as well. The 36 questions are represented by eight subscores that correspond to health concepts, like a limitation in social activities because of physical or emotional problems, or bodily pain, vitality, etc. [Fra92]. The calculation of the subscores is the same as with the subscores from the VF25 questionnaire.

In addition to the subscores, the SF36 questionnaire provides two higher-ordered summary scores called “Mental Component Summary” (MCS) and “Physical Component Summary” (PCS). All of the 8 subscores contribute to one of those summary scores, for example, the scores Vitality, Social Functioning, Role-Emotional, and Mental Health are connected to mental health and, therefore, are used for the calculation of the MCS [War01]. The other four subscores contribute to the PCS.

The calculation of the two summary scores is different from the one from the subscores. It consists of three steps where the first one is the standardization of the subscores “using means and standard deviations from the general U.S. population” [WMK93, p. 4:1].

Then the standardized scores are used for the calculation of the aggregated PCS and MCS based on factor score coefficients again from the general U.S. population. Finally, the PCS and MCS are transformed to norm-based scoring with a mean of 50 and 10 as standard deviation [WMK93].

Those two questionnaires are the main source of the domain knowledge that will be included in the approach. When talking about domain knowledge in the context of KAVA it needs to be distinguished between declarative and procedural knowledge. Declarative knowledge, also called descriptive knowledge, focuses more on the data and descriptions of it and the procedural more on procedures and their best practices [FWR<sup>+</sup>17]. In our case, the questionnaire data, as well as the subscores, are declarative knowledge. Domain experts developed the questionnaires and defined questions that can represent the quality of life of the patients. Furthermore, it was defined which categories of quality of life are covered by the questionnaires and which questions contribute to which category which is also declarative knowledge.

### User

The users of the approach are expected to have domain knowledge. Therefore, possible users are health professionals, like physicians, nurses, or medical researchers, possibly with a special interest in Uveitis. Furthermore, domain experts that are involved in managing and analyzing clinical trials are possible users as well.

### Task

The main tasks of the users of our approach are:

1. get a general overview of the data from the clinical trial
2. explore the patient characteristics as well as the questionnaire results
3. explore how the quality of life evolved over the visits
4. compare the quality of life of certain groups of patients which each other, i.e. if differences between the treatment groups exist

5. find patterns or associations between certain factors
6. get more details on demand with interaction and exploring techniques, like highlighting, filtering, etc.

### 5.1.2 Nested Model by Tamara Munzner

The second model we used for the problem characterization is the “Nested Model” from Munzner [Mun09] which was already presented in Chapter 2 in Figure 2.1. The basic idea behind this model is to guide the design process of visualization systems by explicitly defining the different stages that a design process should consist of and the order in which they should be conducted. In the nested model the output from one level is the input for the next level. Furthermore, the model suggests validation approaches for each level. “Each level in this model has a different set of threats to validity, and thus requires a different approach to validation” [Mun09, p. 922].

For example, if in the first level the wrong problem is defined it will lead to a visualization system that does not fit the actual problem of the target users or a wrong data abstraction might lead to a visualization that conveys the wrong information [Mun09]. To avoid such problems we decided to use this model as a base for our design process phase.

#### *Level 1 - “domain problem characterization”*

The first level of the model is called “domain problem characterization” which we already finished in the step before with the help of the design triangle. Therefore, we were able to already start with the next level “data/operation abstraction design”.

#### *Level 2 - “data/operation abstraction design”*

At this level, the characterization of the problems and the data from the level above are mapped to more abstract operations and data types in the vocabulary of the information visualization domain [Mun09]. As already discussed our dataset consists of different data types which we defined in detail in this step.

The dataset, in general, is multivariate since it resulted from a clinical trial with 190 patients. The personal data from the patients is a combination of qualitative as well as quantitative data. The qualitative data consists of nominal data, like gender, country, Uveitis classification, etiology, etc. as well as ordinal data, like the age categories. Furthermore, quantitative data, like weight, height, body-mass-index (BMI), and visual acuity are included.

The data from the questionnaires is again ordinal but considering that the questionnaire data was collected multiple times it is time-oriented data as well. In addition to the questionnaire data, the subscore data is available as well which is quantitative data and again time-oriented because the subscores were calculated for each of the three visits. The last part of the data is the aggregated data we calculated by ourselves which is ordinal data since it is a value ranging from -2 to +2 for each patient and each question.

### *Level 3 - “encoding/interaction technique design”*

Based on the results from the second level the visual encodings and interaction of the visualization system were designed which is the third level “encoding/interaction technique design”. For this level, the findings and experiences from the exploration phase with Tableau were used, and based on these findings mockups were created which are presented and explained in detail in the next Section 5.2.

### *Level 4 - “algorithm design”*

The last level “algorithm design” deals with the implementation of a prototype based on the final mockups from level 3. The process, as well as the results from this step, are presented in Chapter 6.

## 5.2 Mockups

As already mentioned, in this section the process of the mockup development is explained in detail. The mockups were regularly discussed with visualization experts as well as an domain expert and improved based on their feedback.

### **Encoding and Interaction Technique Brainstorming**

The mockup development is part of the third level of the nested model. Therefore, the first step in this process was the design of encodings and interaction techniques based on the output from the second level as well as the results from the data exploration phase with Tableau. When exploring the data with Tableau multiple encodings and interaction techniques for the different parts of the data were already defined. Therefore, we needed to decide which of the encodings are suitable for the final approach or if new ones are needed.

We started with the questionnaire data where multiple encodings, like the heat maps or the sparklines were already tested. However, since the heat map was the most space-efficient encoding and matrix-like visualizations, like heat maps can reveal patterns in aggregated data the decision to use the heat map was made [GES16; LSA<sup>+</sup>16]. One feature from the heat maps in Tableau that was retained as well are tooltips that provide details about the patients on demand.

For the subscore data, bar charts and radar charts were used in Tableau. An advantage of the radar chart is that each line in the chart has a certain shape which makes comparison easier than in the bar chart. It is possible to check if shapes are similar or if some of them have overlaps which are especially helpful for the detection of a temporal evolution from visit to visit. Furthermore, the radar chart is more space-efficient than the bar chart because the dimensions always remain the same regardless of which characteristic is compared in the chart because only the number of lines changes, and in the bar chart more bars are added which results in a bigger chart width. These were the main reasons for the decision to use the radar chart for the subscore data visualization.

The last part of the dataset is the patient data where visualizations, like a world map or bar charts were used in Tableau. However, since the main focus of the approach is not on the visualization of the patient data itself and the world map for example only conveys very limited information about the patients and still needs a lot of space we decided to not include it in our approach.

Additionally, we decided that our approach should include filters, like in the Tableau dashboards.

### Mockups - Version 1

After the initial encodings and interaction techniques were decided the mockup development started by using the wireframing software Balsamiq<sup>1</sup> in combination with screenshots of the charts created in Tableau.

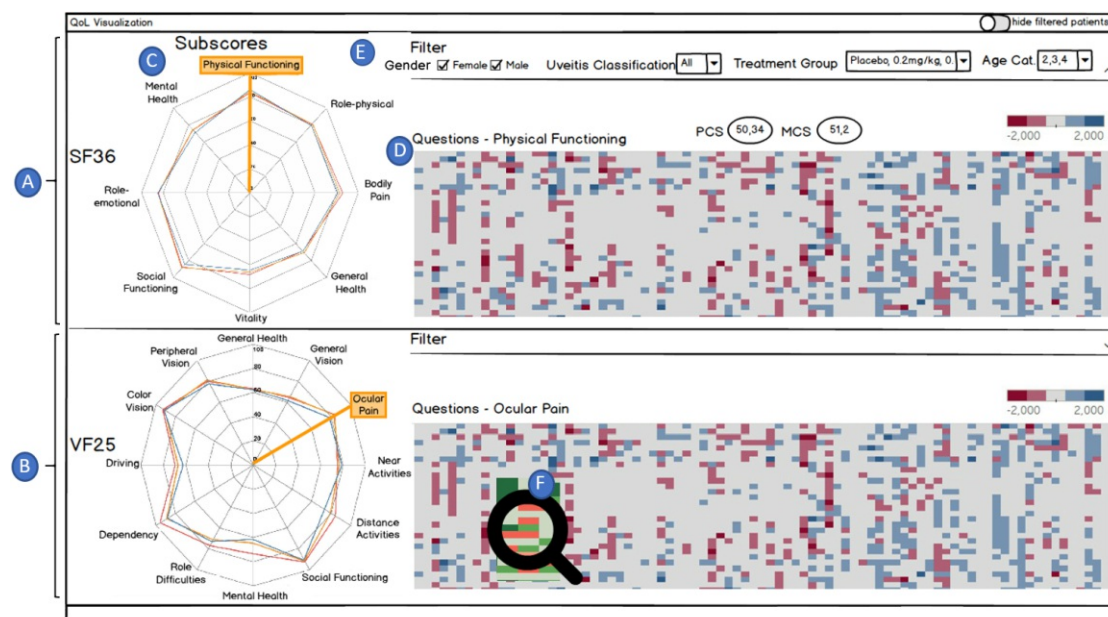


Figure 5.1: First mockup of own approach: The view is vertically split where (A) includes visualizations of the SF36 questionnaire data and (B) visualizations of the VF25 questionnaire data. Both halves include the same visualizations, firstly, the radar chart (C) visualizing how the average subscores have changed from visit to visit, and, secondly the heatmap (D) visualizing how the answers from each patient for each question have changed over time. (E) is a filter pane that includes multiple filters that affect both, the radar chart as well as the heatmap. (F) illustrates the lens interaction technique that reveals the data behind the aggregated values which are the answers of the patients for the questions at all visits.

For the brainstorming of possible mockups, the tasks from Section 5.1.1 were taken into

<sup>1</sup><https://balsamiq.com/> (accessed 30/04/2021)

account to find an approach that can support those tasks properly. The first consideration made was whether to show the data from both questionnaires at once or to add the ability to choose one of the two. The advantage of choosing the questionnaire is that more space is available because the view does not need to be split in half. However, in the end, it was decided to show both questionnaires at once because it makes a comparison between the two questionnaires possible, and since there are some parallels between them this could be helpful. Therefore, it was clear that the view needs to be split in half. The next step was to try out how the visualizations, like heat maps and radar charts could be arranged properly and since the heat map is the visualization that needs the most space we designed the view around the heat map.

A first mockup evolved as can be seen in Figure 5.1. The view was vertically split and on top (A) the data from the SF36 questionnaire is visualized and below (B) the VF25 questionnaire data. For both questionnaires, the visualizations and filters are the same. On the left, the radar chart (C) visualizes how the subscores have changed on average from visit to visit and on the right, the heat map (D) with the aggregated data is displayed. On top of the visualizations, a filter pane (E) is added which is collapsible as can be seen in the VF25 half. Furthermore, we had the idea that the radar chart could be used as a selection to be able to choose a certain subscore and display only the questions that belong to the subscore in the heat map. The selection is indicated by an orange rectangle and line in the radar chart.

Another interaction technique included in the heat map is a lens that can be seen in (F). The idea behind this lens is that in the heat map only the aggregated data are visualized and it might be interesting to see the exact answers the patient gave at the different visits for a certain question to see how the aggregated value was calculated. Therefore, the idea was that on hovering over the heat map a lens is overlaid that shows the answers from the questions per visit. In the mockup again a heat map visualization is used inside the lens but we also had the idea of maybe showing another visualization, like a line chart or sparkline. The idea of the lens is based on the grouping lens from Gnaeus [FUA<sup>+</sup>15].

Another feature introduced in this mockup is the switch in the top right corner where it can be selected whether patients or questions that get filtered out are removed from the view or only greyed out. The idea behind this feature is that it might be interesting to have the filtered patients/questions as a reference and, therefore, should not be removed but the advantage of removing them is that the data gets easier to read because of the newly generated space. This first mockup was discussed in a meeting with visualization experts. In general, the mockup was received very well but some suggestions for improvement were provided. For example, to separate the rows in the heat map by subscore using lines to be able to see what questions belong to which subscore. Another feature that was suggested is to add the ability to sort the heat map, for example, to be able to order the patients by certain characteristics, like age or gender which could help to find patterns. It was also suggested that it might be helpful for pattern identification if the user could swap the questions or patients on their own.

The final feedback was regarding the lens which is a feature they liked but also suggested



to think about the concrete functionality in more detail. For example, if it should show the heat map with the detailed answers as in the first mockup or if a tooltip with a line chart would be more effective, etc.

The last hint from the visualization experts was to take a look at the tasks from the paper from Preim and Lawonn [PL20]. They present typical VA tasks, like exploration, comparison, association, etc. that should be covered from a VA approach. These tasks from the paper can be helpful to check whether the mockup covers those or if some of them could be covered by adding or changing certain features.

## Mockups - Version 2

Based on the previous feedback the mockup was adapted and improved and the results of this process are described and presented in this section. The first step was to incorporate the feedback from the previous meeting, like the separation of questions based on subscore in the heat map which can be seen in (A) in Figure 5.2 where the grey lines indicate which questions belong to which subscore.

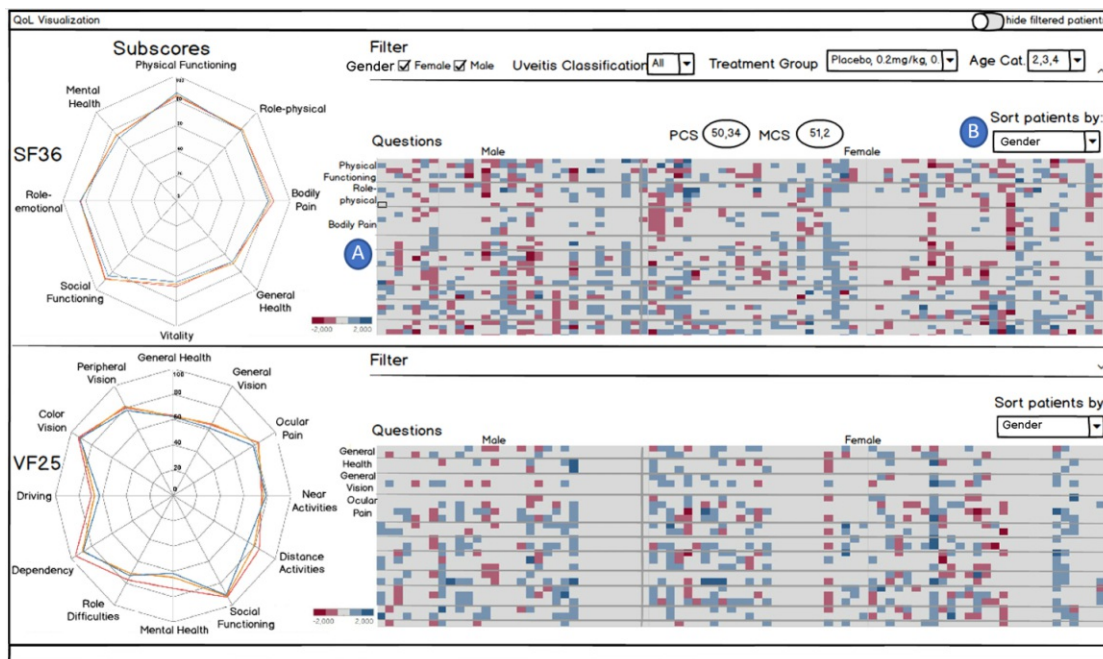


Figure 5.2: Updated mockup: (A) The rows in the heat map are separated using lines to indicate which questions belong to which subscore. (B) A dropdown input that enables sorting of the heat map by certain characteristics, like gender, Uveitis classification, etc.

Furthermore, a dropdown input field was added on top of the heat map (B) that enables the user to sort the heat map by certain characteristics. For example, in Figure 5.2 it was selected to sort the patients in the heat map by gender, and a vertical separating line is added as well as labels that indicate the sorting.

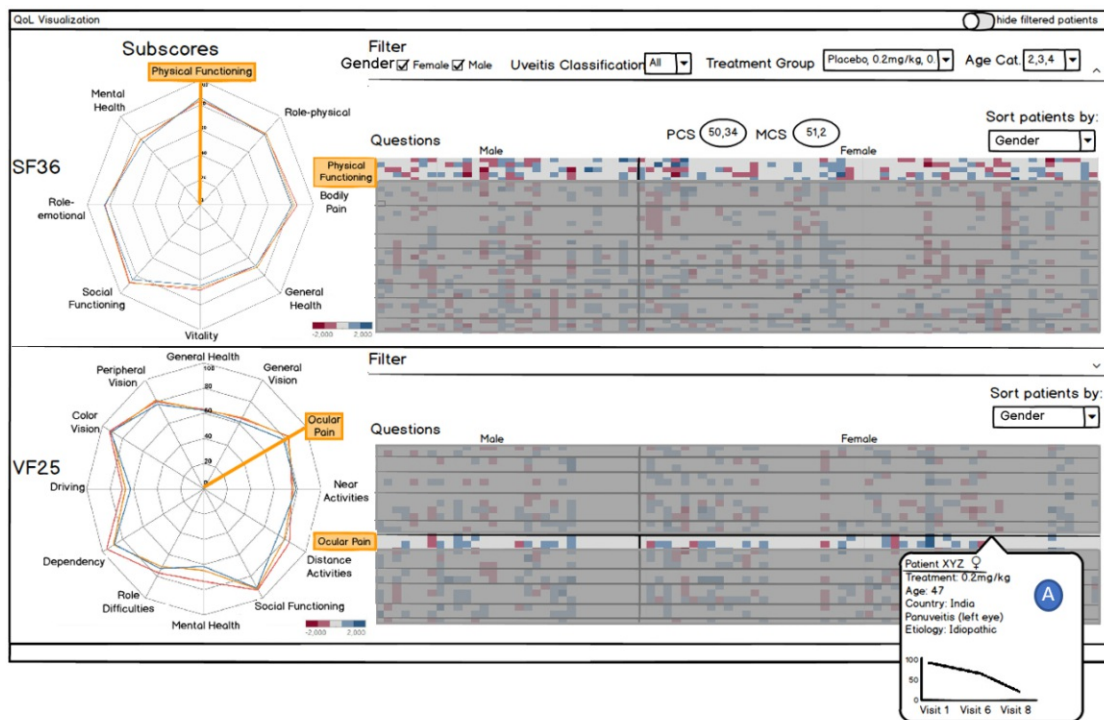


Figure 5.3: Updated mockup with details on demand: (A) shows the tooltip that is displayed when hovering over a rectangle containing patient-related data, like age, gender, country, etc. as well as a line chart visualizing how the patient answered the selected question at each visit. In the radar chart certain subcores were selected which are highlighted in the heatmap by greying the others out.

Regarding the feedback for the lens functionality, we decided to replace the lens with a tooltip that displays, on the one hand, patient-related data, like gender, age, Uveitis classification, etc., and, on the other hand, a line chart that displays how the selected patient answered the selected question at each visit. This decision was made because more information can be conveyed with the tooltip than with the lens approach where only the answers from the question would have been displayed. Another reason was that with the previous solution patient-related data was not incorporated in the approach and since this information is important as well the tooltip was the better solution. An example of this tooltip can be seen in (A) in Figure 5.3. In this mockup, it can also be seen that certain subcores were selected in the radar chart and the corresponding rows in the heatmap are highlighted by greying the other questions out.

As already discussed, we provided the possibility of being able to select whether filtered out data should be removed or greyed out. For the case that patients or questions are removed when filtering, we had the idea that if enough space is available the heatmap could be swapped with, for example, a sparkline chart to see at a glance how the answers



have changed from visit to visit. A possible version of this feature can be seen in (A) in Figure 5.4.

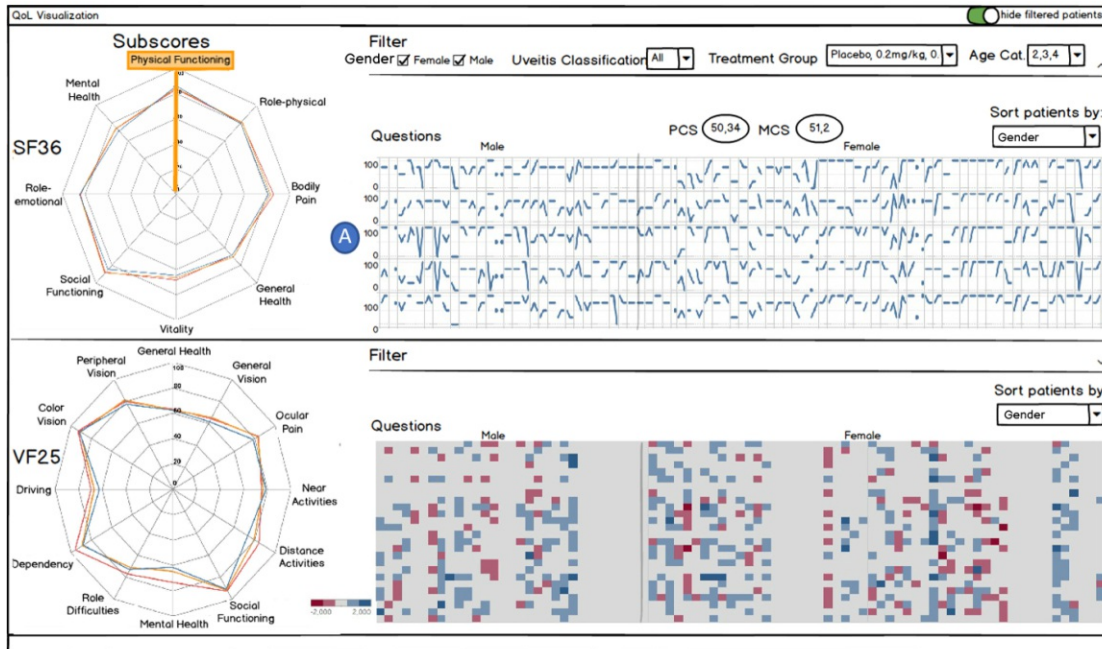


Figure 5.4: Updated mockup with a new feature: (A) shows sparklines that visualize how each patient answered each question at the different visit which is only displayed if enough space is given through filtering.

After including the feedback from the previous meeting the second step was to go through the suggested paper from Preim and Lawonn [PL20] and analyze how the tasks from their paper are incorporated in our approach and where improvement is possible. Only tasks T1 to T5 were analyzed because the others are more specific for VA approaches of public health data which is not the case in our approach.

However, the first task that is presented in their paper is “T1 Exploration” whose goal is that VA approaches should assist in exploring the data and gathering information about it [PL20]. In our opinion the adapted mockups fulfill this requirement because the approach enables the user, on the one hand, to explore the questionnaire data in an aggregated form that informs the user whether the quality of life has improved or worsened but also in the raw version, and, on the other hand, the temporal evolution of the subscores can be analyzed as well as information about the patients is given.

The second task “T2 Assessment and pattern identification” is covered in our mockups with the filters as well as the sorting and reordering functionality of the heatmap but greater support could be provided.

“T3 Associations” means to be able to find associations between certain factors, like

environmental factors, type of disease, etc. [PL20]. By providing the possibility of filtering and sorting, associations between certain variables could be revealed with our VA approach. Nevertheless, this task could be more simplified and, therefore, we decided to keep this aspect in mind and try to find a better solution.

The next task “T4 Verification” is connected to the previous task because its goal is to provide the possibility of verifying certain associations or patterns and this is again only possible through the filtering and sorting features in our approach.

“T5 Comparisons” is the last task considered from the paper and the intent behind this task is that a VA approach should enable the user to make comparisons between certain groups of patients which is possible by using the sorting functionality. However, since the sorting functionality only affects the heat map and comparison in a visualization with such an amount of data is more difficult than it would be, for example, by using the radar chart it was decided that this task is an aspect that should be given more attention during the future mockup development as well.

All in all, many of the tasks were already covered by our mockups but in some areas, there was still room for improvement which is the reason why we decided to try to find another approach that is different from the first one.

In the next meeting with the visualization experts the adapted mockups (Figure 5.2, Figure 5.3 and Figure 5.4) from the first version were discussed.

Some feedback we received was to add the ability to see the question texts, for example, by adding an icon to the left of each row and on hovering over this icon the full question text gets displayed. Furthermore, axis indicators for the heat map were missing. The experts liked the new sorting/ordering feature but suggested that a sorting functionality for each question could also be helpful.

Regarding the idea with the sparklines the feedback was given that the lines could be colored the same as the rectangles in the heat map to have the encoding of improvement or worsening there too.

### Mockups - Version 3

After the second meeting again the feedback was incorporated into the mockups and the result can be seen in Figure 5.5. The main differences to the previous version are that all subscores are labeled in the heat map (A) and to the left of each row an “information” icon was added, as discussed with the experts, that shows the whole question text on hovering (B). An interaction technique that was defined for the subscore labels is that it is possible to reorder columns per subscore by clicking on a subscore label and dragging it above or below the other subscores. This feature might be helpful for better comparison between certain subscores or for pattern identification.

Furthermore, the axis indicators were added as well as the “[shuffled]” keyword (D) to indicate that the data were shuffled due to sorting or grouping. The last addition is that a new column was added at the end of the heat map that presents the average value of

each row. The idea behind this column is that it might be interesting to see if certain questions improved more or less than others considering the average of all patients.

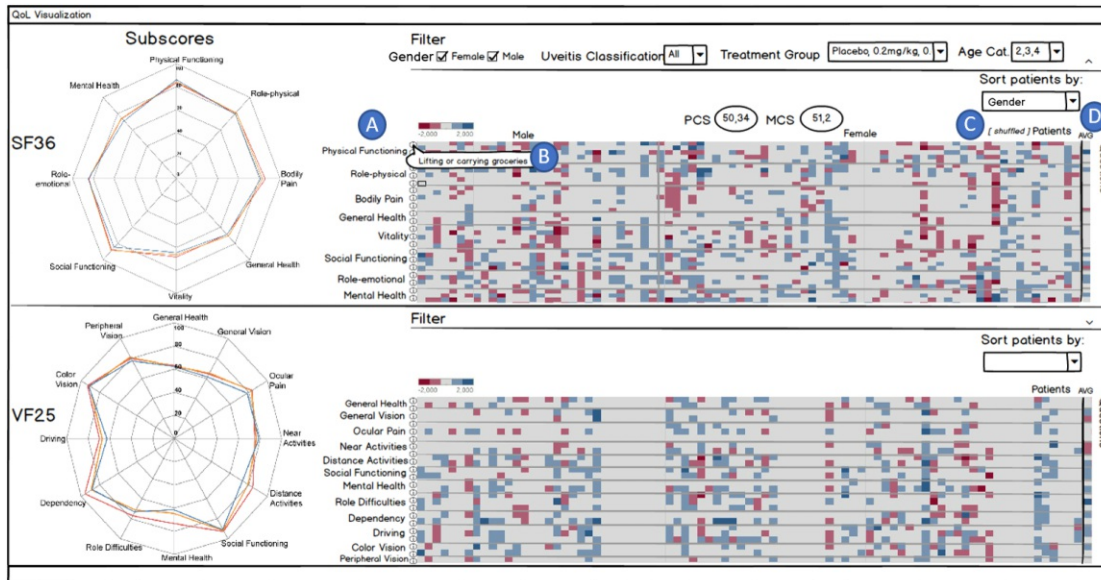


Figure 5.5: Updated mockup: All subscore labels were added to the questions (A), (B) shows a tooltip with the full question text that is displayed on hovering over the “i” icon. (C) Axis indicators were added to the heat map as well as the “[shuffled]” indicator and (D) is a new column at the end of the heat map visualizing the average value over all patients for each question.

While finishing the incorporation of the feedback we still brainstormed for another approach that might be able to support some of the tasks from the paper from Preim and Lawonn [PL20] better than the current solution. One aspect that was missing in the current solution was to support the pattern identification and association assessment better. With the heat map and its features, like sorting and filtering it is possible to identify patterns but it is not that easy especially if a person does not know beforehand what type of patterns could be possibly revealed. For example, maybe a certain group of patients with similar characteristics has similar results in quality of life which could indicate a pattern but with the current solution, the person would need to try out the different filters to find groupings which might take more time. Therefore, the main focus of the second approach is more on revealing patterns or associations between certain characteristics and the quality of life parameters.

Since the visualizations themselves already convey all main aspects of the dataset we thought about a different filtering method as an input for the heat map and the radar chart. After some brainstorming and research, the visualization called “parallel set” was discovered which might be a good solution for the filtering. This type of visualization is especially helpful for categorical data which fits the dataset very well since the

characteristics of the patients are mainly categorical, like for example the Uveitis type, the age category, etc. [BKH05].

“Parallel Sets is a new visualization method that adopts the layout of parallel coordinates, but substitutes the individual data points by a frequency-based representation” [BKH05, p. 133]. An example can be seen in (A) in Figure 5.6. The basic idea behind the parallel set visualization is that instead of using a continuous scale, like in a parallel coordinates chart, sets of boxes representing the different categories are used whose size depends on the frequency of the category [BKH05]. Furthermore, since the layout is similar to the one of a parallel coordinates chart each category has its own axis which makes it possible to see how the different categories are split and which connections they have to each other. Those aspects make it a good visualization for the patient data because with the parallel set it can be analyzed how the different patients are distributed among the different categories.

Another advantage of the parallel set visualization is that it comes with different interaction techniques, like the possibility of reordering the dimensions or on hovering over a box a tooltip shows the exact frequency of the category. Furthermore, on hovering over a box the box itself and all boxes that are connected to this box are highlighted as well which makes it possible to identify even better how the different categories are connected.

Therefore, the decision was made to use the parallel set visualization for the alternative approach. The idea was that it can be used, on the one hand, to explore the patient data, to see how they are distributed, and, on the other hand, to use it as a filter for the other visualizations of the view. As already mentioned, the other visualizations, like the heat map and radar chart fit very well to the dataset which is the reason why we tried to stick with those visualizations if possible and combine them with the parallel set. However, since the first approach did not have enough space for the parallel set visualization the components needed to be rearranged.

The result of the rearrangement and the final mockup can be seen in Figure 5.6. The main idea behind this approach is that there are two different options to filter the patients that should be displayed in the other charts which we called “Cluster-based” patient selection and “Filter-based” selection. The “Cluster-based” selection is the one using the parallel set which can be seen in (A) in Figure 5.6 and the other one is the same selection as in the previous mockups by using simple dropdown inputs or checkboxes as filters as can be seen in (B) in Figure 5.7. In the first mockup in Figure 5.6 on the left the parallel set chart (A) with some example categories, like the Uveitis classification and the gender is illustrated. By clicking on one of the boxes the patients belonging to this box are selected and visualized in the other visualizations on the right. Below this chart in (B) is the second type of selection which can be expanded by clicking on the label and then different dropdown fields will be shown which can be seen in (B) in Figure 5.7.

In (C) some general statistics regarding the dataset are provided, for example, the number of patients and how many of those are male or female or how the different Uveitis classifications are distributed. These statistics are only some placeholders and could be



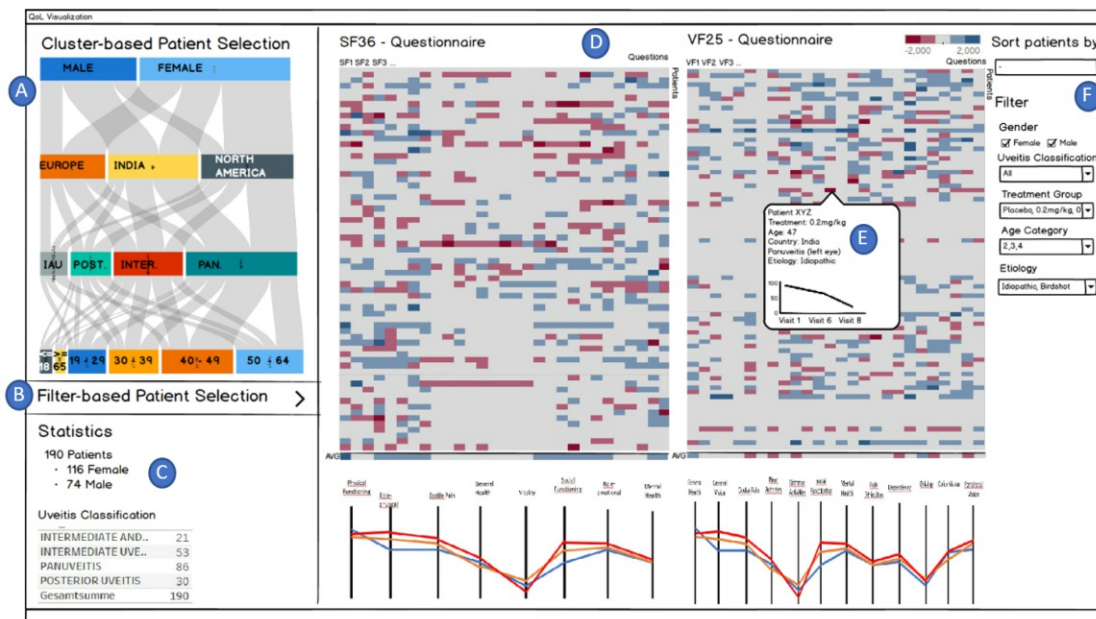


Figure 5.6: New approach - “Cluster-based” selection: (A) Parallel set visualizing how the patients are distributed among categories, like gender, region, or Uveitis classification which can be used as a filter for the other charts by clicking on one of the boxes. (B) is the currently collapsed “filter-based” selection and (C) provides some general statistics about the data. (D) is the main view that is split in half for the two questionnaires where each of them consists of a heat map and a parallel coordinates chart. (E) displays a tooltip providing patient-related data as well as data about the questionnaire and (F) is a filter pane with some filters as well as the sorting feature.

replaced by other statistics in the final approach if necessary.

The center part of the screen (D) is again split in half but this time vertically and not horizontally since it was not possible otherwise. Therefore, the heat map axes were swapped and the x-axis represents now the questions and the y-axis the patients because in this way the heat map has a smaller width since the number of patients is higher than the number of questions. Below the heat map, a parallel coordinates chart is used instead of the radar chart. The main reason for this change is that the radar chart needs more space regarding the height compared to the parallel coordinates chart and because the heat map is already very space-consuming especially vertically a visualization that is wider but not as tall was needed. Furthermore, since the only difference between those two visualizations is the arrangement of the axis we decided that this would be the best solution.

Another advantage of the parallel coordinates chart in combination with the heat map on top is that in the parallel coordinates chart each axis represents one subscore and in the heat map the x-axis represents the questions which makes it possible to align the

two charts by their axis by ordering the questions in the heat map in a way that the questions belonging to the subscore in the parallel coordinates chart are aligned.

As in the other mockups, this approach again provides tooltips (E) on hovering over a rectangle in the heat map that presents patient-related information and a line-chart visualizing the exact answers from the selected question for each visit.

The last part on the right (F) includes the same sorting feature as in the previous mockups which sorts the heat map by a certain characteristic, like gender or Uveitis classification. Furthermore, some additional filters are included in this side pane as well.

As it can be seen, many of the features from the previous mockups stayed the same, like the sorting feature or the tooltips. Furthermore, something that is not shown in the mockup but is still part of it is that the question text would be shown on hovering over the question or that the heat map can be sorted by question by clicking on one of the question labels.

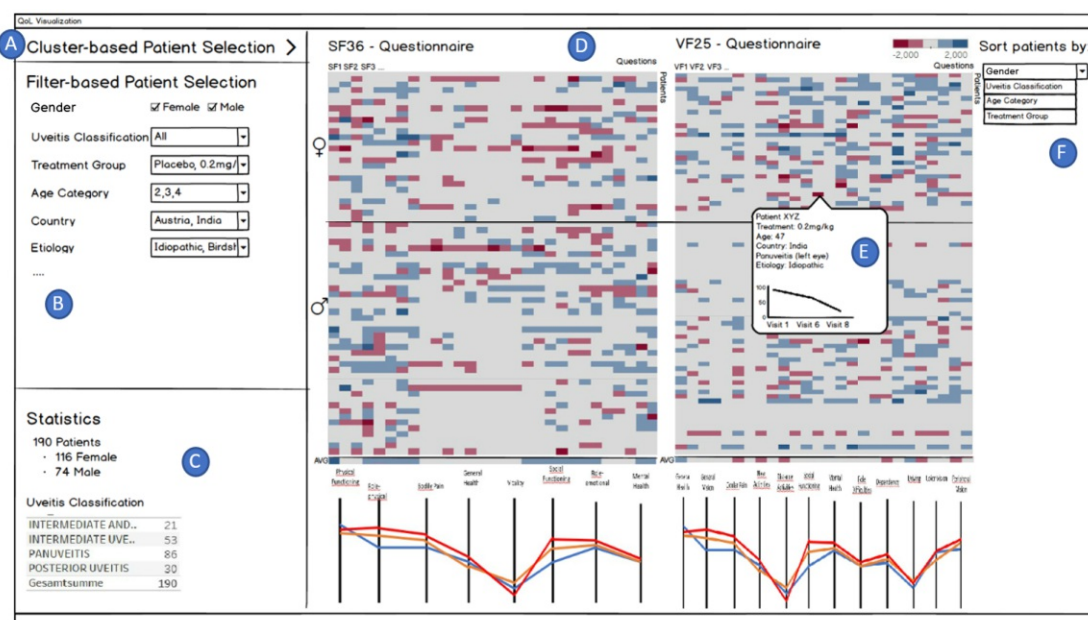


Figure 5.7: New approach - “Filter-based” selection: (A) indicates the currently collapsed “cluster-based” selection. (B) is the “filter-based” selection with some dropdown inputs and checkboxes that are responsible for which patients are displayed in the other charts. (C) is again the area with some general statistics about the dataset. The main part (D) is the same as in the other part with the heat map and parallel coordinates chart for each questionnaire as well as the tooltip (E). With the dropdown input (F) it can be selected how the patients should be sorted in the heat map which is in this case by gender which is also displayed in the heat map.

As already mentioned, there are two different types of filters and the second one can be seen in Figure 5.7. (A) is the “Cluster-based” selection part from the previous mockups

which is collapsed in this screen and in (B) the “Filter-based” selection can be seen. This type of selection includes the same dropdown inputs and checkboxes as from the previous mockups which can be used to filter what patients should be visualized in the charts on the right. (C) to (E) are the same as in the first mockup with the same functionalities. The only difference is that in the heat map the sorting feature is illustrated. (F) is again a filter pane but only with the ability to sort the heat map by a certain characteristic since the other filters are already included in (B).

After these mockups were finished again a meeting with visualization experts was conducted. Regarding the first mockup, no further changes or suggestions were given. During this meeting, the new approach was presented as well and they were satisfied and liked the new filtering approach with the parallel set. Some feedback that was given was that it might be good to include the percentages into the boxes of the parallel set if enough space is available.

Regarding the two different types of selection, they had the idea of removing the “Filter-based” selection and instead make the parallel set configurable to be able to select what categories should be displayed in the parallel set. Furthermore, for even better comparison abilities it was suggested to investigate if multiple selections in the parallel set would be possible and how it could be incorporated in the other charts. The last feedback was that as with the other mockups it would be good to have the subscore labels on top of the heat map and the same reordering functionality.

All in all, the new mockup was received very well and it was suggested to stay with this approach and that it is already in a stage where a prototypical implementation could be started. Therefore, the next step was the planning and development of the prototype whose entire process is presented in Chapter 6. During this process, we continued with the regular meetings where constant feedback and improvement suggestions were given from the visualization experts as well as the domain expert.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.



# Prototype Implementation

After the design process was finished a prototype was developed based on the mockups. In this chapter, the technologies used as well as the final prototype including screenshots is presented. The prototype is available online under <https://knova-app.herokuapp.com/> (accessed 18/05/2021).

## 6.1 Technologies

For the implementation of the prototype the following frameworks and libraries were used:

- Angular
- D3.js
- d3.parsets
- Heroku

### Angular

Angular<sup>1</sup> is a front-end web application framework based on TypeScript<sup>2</sup>. The framework is component-based and is able to build scalable web applications ranging from small projects to enterprise-level applications. Angular has multiple advantages, like the existence of a variety of libraries that make the framework even more powerful, for example, libraries for the management of forms or client-server communication. Furthermore, Angular provides developer tools that make developing, building, and testing easier. Angular was used as the main framework for the prototype.

<sup>1</sup><https://angular.io/> (accessed 29/04/2021)

<sup>2</sup><https://www.typescriptlang.org/> (accessed 29/04/2021)

### D3.js

D3.js<sup>3</sup>, short for Data-Driven Documents, is a JavaScript library intended for the manipulation of documents based on data. D3.js uses web standards, like HTML, CSS, and SVG to create interactive data visualizations that are fast and able to handle large datasets.

In our prototype D3.js is used for the development of all visualizations, like the heat maps or the parallel coordinate charts.

### d3.parsets

d3.parsets<sup>4</sup> is a reusable parallel set chart built with D3.js. It is configurable and includes multiple interaction techniques. We used d3.parsets for our parallel set visualization and added some additional features, like being able to select a certain set of the chart and highlighting it.

### Heroku

Heroku<sup>5</sup> is a platform as a service (PaaS) that can be used to build, run, and deploy applications in the cloud. We used Heroku for the deployment of our prototype to make it available online which makes it possible to share it more easily with others, for example with participants of user studies or the domain expert.

## 6.2 Graphical User Interface Composition

The graphical user interface of the prototype is vertically split into three views, as can be seen in Figure 6.1. The first view on the left (A) is the patient selection view with a form for selecting categories and a parallel set chart. In the center is the main view (B) including the main visualizations of the data from the two questionnaires and on the right (C) the filter view is present. Furthermore, the header (D) includes a search bar as well as a menu for changing modes which is explained in detail later.

## 6.3 Views and Components

In this section, all views, components, and features of the prototype are presented. Furthermore, explanations for certain deviations from the mockups are given. In the prototype the countries and regions were renamed to Disneyland names which was suggested from the domain expert to avoid any biases during the evaluations.

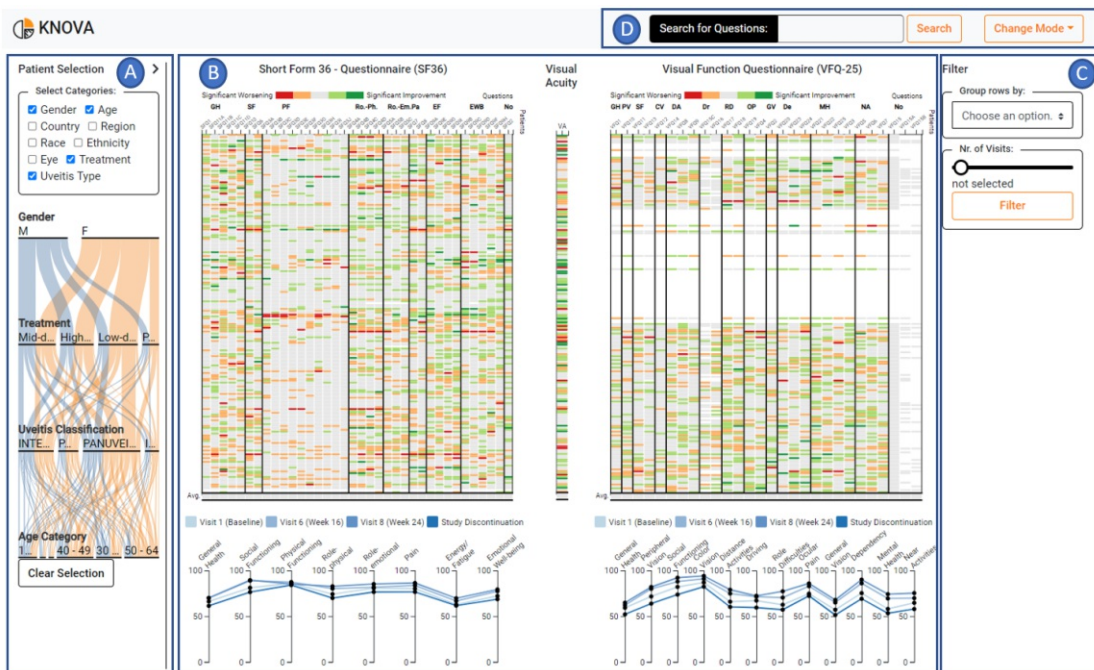


Figure 6.1: Prototype layout: The user interface is horizontally split into three views: (A) the patient selection view, (B) the main view, and (C) the filter view. The header (D) includes a search bar and a menu for changing the mode of the prototype.

### 6.3.1 Patient Selection View

The first view on the left is the patient selection view which can be seen in detail in Figure 6.2. This view consists of a section containing multiple checkboxes (A) as well as the parallel set visualization (B). With the checkboxes, it is possible to select which categories, like gender, age, Uveitis type, etc. should be visualized in the parallel set. When selecting or deselecting one of the checkboxes the parallel set gets updated immediately. Per default, gender, age, treatment, and Uveitis types are selected. As already discussed, the parallel set has two major use cases. The first one is that it visualizes how all the patients are distributed among the different categories. The second one is that it can be used as a filter. By clicking on a category, like male or on one of the strands, like females that received the placebo this group of patients is selected and is used as an input for the charts in the main view where only the selected group of patients is then visualized.

Further interaction techniques included are the possibility of reordering the categories horizontally as well as vertically. By clicking on a category label, like “Gender” and dragging it up or down the order of the categories is swapped. The same can be done

<sup>3</sup><https://d3js.org/> (accessed 29/04/2021)

<sup>4</sup><https://github.com/jasondavies/d3-parsets> (accessed 29/04/2021)

<sup>5</sup><https://www.heroku.com/> (accessed 29/04/2021)

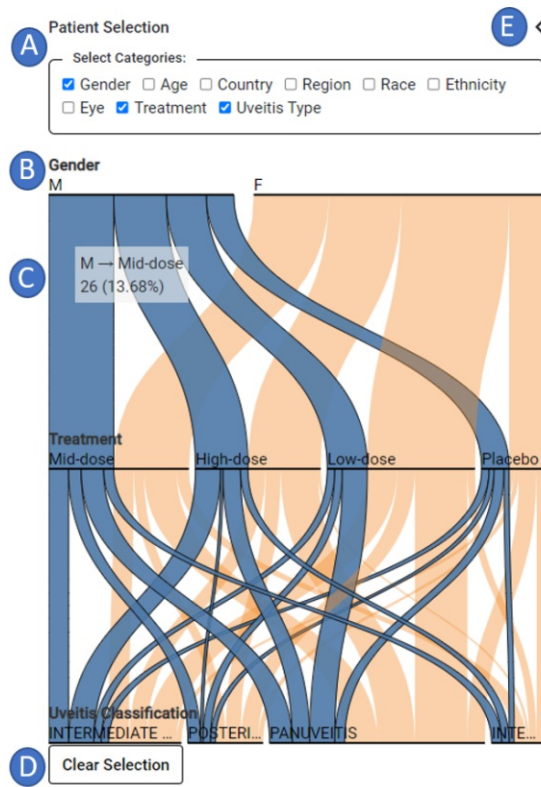


Figure 6.2: Patient selection view: (A) Multiple checkboxes that define which categories are shown in the parallel set (B). (C) is a tooltip showing how many patients belong to the selected set and with the “Clear Selection” button (D) the selection is reset. With the arrow on the top right (E) the patient selection view can be expanded or condensed. In this figure the view is expanded.

horizontally within each category, like swapping the male and female sets. On hovering over a strand or category, like male, the absolute number of patients belonging to this set is shown as well as the percentage which can be seen in (C) in Figure 6.2. In Figure 6.2 it can also be seen that the user selected to visualize only male patients which is indicated by the black border around the strands as well as them having a higher opacity. With the “Clear Selection” button (D) it is possible to clear the entire selection which updates all charts immediately to show all patients again.

The final feature of this view was added after the first evaluation because it turned out that for multiple tasks that deal with figuring out how the patients are distributed between the categories only the parallel set visualization is needed and since it is rather small on the left it was sometimes hard to select especially very thin sets that contain only a few number of patients. Therefore, the possibility of expanding the patient selection view was provided which can be seen again in Figure 6.2. In this Figure, the view was expanded and due to the bigger size, the selection of small sets is easier. The view can

be expanded or condensed to normal size by clicking on the small arrow in the top right corner (E).

When comparing the final patient selection view with the mockups only one part has changed which is the section containing general statistics about the data which was not implemented in the final prototype. The main reason for this change is that there was not enough space anymore and many of the statistics, like how many patients were in the different treatment groups could be figured out with the parallel set as well. Therefore, there was no need for those statistics anymore.

### 6.3.2 Main View

The main view in the center of the prototype has the goal of visualizing the data related to the two questionnaires and mainly stayed the same as defined in the final mockups. Therefore, for each questionnaire a heat map as well as a parallel coordinates chart is included. As already mentioned, the heat map visualizes how the answer of each patient for each question has changed from the first to the last visit ranging from significant worsening to significant improvement indicated with five steps. The parallel coordinate charts are visualizing how the subscore data from both questionnaires has changed on average for each of the different visits.

An additional heat map was added between the two heat maps that visualizes how the visual acuity score has changed from the first visit to the last visit with the same encoding as the other two heat maps as can be seen in (A) in Figure 6.3. This heat map was added because the domain expert saw a need for being able to analyze the change of the visual acuity as well. From the domain expert, we also got the domain knowledge of how the encoding of the worsening/improvement should be calculated. For each patient the visual acuity score from the last visit is subtracted from the first visit and the resulting difference is mapped to colors ranging from red to green. The mapping scheme is shown in Table 6.1.

Difference Range	Color
$x < -10$	red
$-10 \leq x < -2$	orange
$-2 \leq x \leq +2$	grey
$+2 < x \leq +10$	light green
$x > +10$	green

Table 6.1: Mapping scheme of the visual acuity difference value to the color scale of the heat map ranging from red to green.

As already discussed, the heat maps come with multiple interaction techniques. The first one is a tooltip that is shown on hovering over an element in the heat map (see (B) in Figure 6.3). The tooltip contains general information about the selected patient as well



## 6. PROTOTYPE IMPLEMENTATION

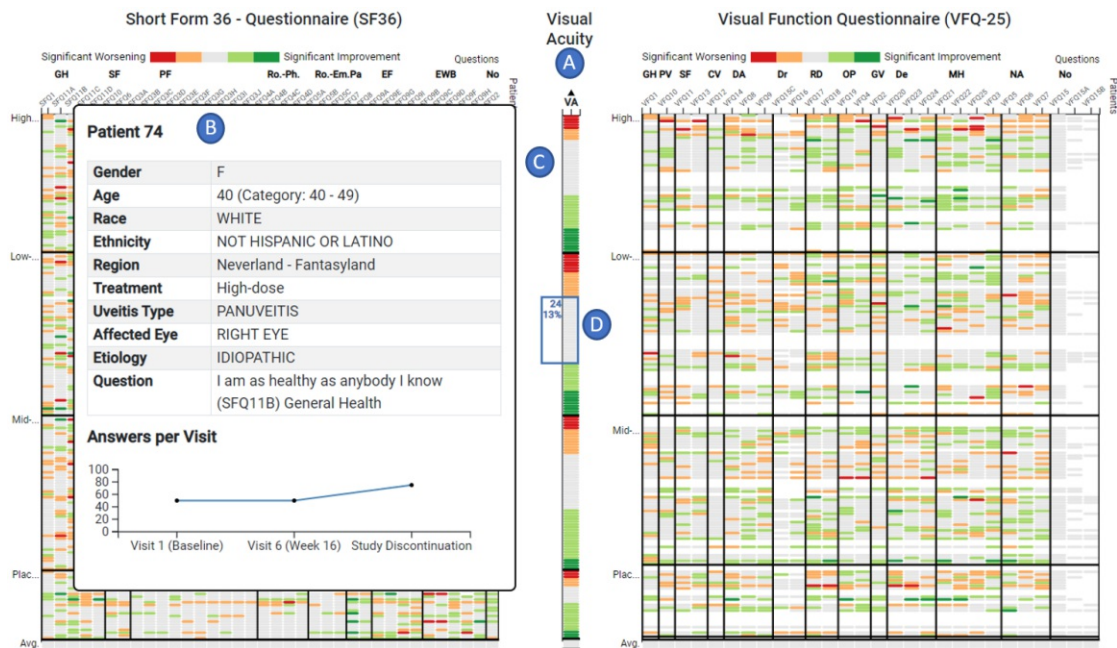


Figure 6.3: Main View - Heat maps: (A) displays the newly added heat map for the visual acuity scores using the same encoding as the other heat maps. (B) is a tooltip containing patient-related information as well as a line chart visualizing the concrete answers per visit. (C) shows that the visual acuity heat map was sorted and since the grouping is active the sorting is done for each group separately. In (D) the counting feature is depicted.

as the question text from the selected question. Furthermore, the exact answer values for the question at all visits are visualized in a line chart.

Another feature is that the heat maps can be sorted by clicking on the label of a column. An example can be seen in (C) where the visual acuity heat map was sorted. The sorting of one heat map affects all other heat maps as well and if the heat map is grouped as it is in the figure then each group is sorted separately.

Furthermore, there is the possibility of grouping the rows of the heat map by certain categories, like treatment groups, gender, etc. which is also shown in Figure 6.3 where the heat maps are grouped by treatment groups indicated by horizontal separating lines as well as labels next to the heat map.

It is also possible to reorder the columns per subscore by clicking on the subscore labels on top of the heat map and dragging them to the left or right to swap positions.

One interaction technique that was additionally included is a counting feature which can be seen in (D) in Figure 6.3. With this feature, multiple rows can be selected in the heat map and the selected rows are counted. The count as well as the percentage value is displayed next to the selected rows. The percentage is calculated using the current total

number of patients that are displayed in the charts meaning that if filters are selected these are considered for the calculation.

The selection of the rows is done with two clicks on any row of the heat map. The first click indicates at which row the counting should start and the second click where it should end and all the rows between the selected are then counted. A right-click on the heat map clears the selection. This feature was a suggestion from one of the visualization experts because for some tasks it is helpful to be able to count the number of rows to for example get to know how many people had a significant worsening or improvement in one of the questions.

In general, the basic features of the heat map were kept as they were defined in the mockups. The only difference is that one additional heat map for visual acuity was added and the color scheme has changed since the domain expert suggested using red and green to encode worsening and improvement.

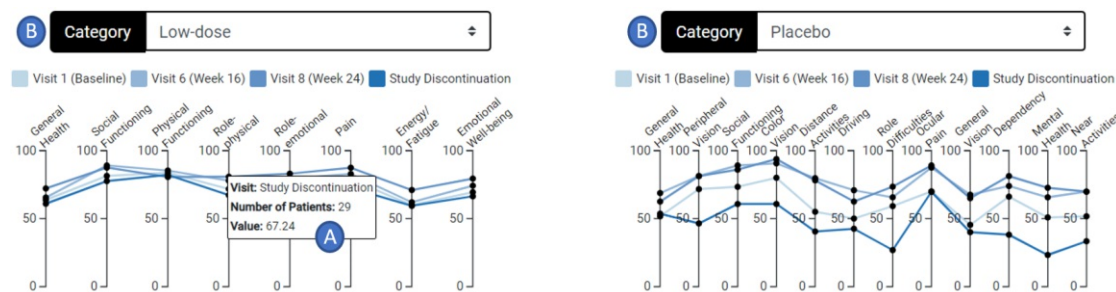


Figure 6.4: Main View - Parallel coordinate charts: (A) shows the tooltip that is displayed on hovering over one of the dots providing information about the visit type as well as the actual value and the number of patients that were used to calculate the average value. (B) shows dropdown inputs that are shown if the grouping filter is active. It enables the user to switch between certain groups of patients, for example, patients from the placebo group and the low-dose treatment group.

The second part of this view are the parallel coordinate charts which can be seen in detail in Figure 6.4. The structure of the parallel coordinate charts stayed the same as defined in the mockups. One difference is that the colors of the lines were changed to a sequential color scheme. The reason for this change is that the colors represent the different visits that have a certain order and with a sequential color scheme this order is better represented.

Furthermore, some additional features were added as well. The first one is the tooltip (see (A) in Figure 6.4) that is displayed on hovering over the circles. The tooltip gives information about which visit this line represents as well as the number of patients that were used to calculate the average score and the actual average value. The number of patients was added since the domain expert informed us that it is important to know how many patients were used to calculate the average score. The reason for this is that, for example, when filtering for patients from the placebo group the majority of those

attended only the first visit and dropped out after that, and only a very few patients of the placebo group completed the study and came to the last visit. Therefore, it is possible that in some cases a line represents only one patient because the others did not attend this visit and this could lead to a misconception.

Another additional feature that was added to the parallel coordinate charts are the dropdown inputs on top of the charts (B). Those dropdown inputs are only visible if the grouping filter is active. When activating the grouping by a certain category, like treatment group there is the possibility to select the different treatment groups in the dropdown input which is then used as a filter for the parallel coordinates chart. For example, as can be seen in Figure 6.4 in the left dropdown input the “Low-dose” treatment group was selected which means that in the parallel coordinates chart only the patients belonging to this group were used for the calculation of the average scores and the same applies in the right chart for the placebo group. This was again a suggestion from the domain expert since it was emphasized that comparing groups of patients is an important task. Especially the comparison of the average subscore values is very helpful and this is the reason why this feature was added. However, one problem with this solution for comparing groups of patients is that the base for the comparison is not the same because the two parallel coordinate charts show different data from different questionnaires and, therefore, another solution was needed as well. To support such comparison tasks even better another functionality was implemented which is explained in detail in Section 6.3.5.

### 6.3.3 Filter View

The filter view contains two different functionalities as can be seen in Figure 6.5. The first one is a dropdown input (A) where the category by which the heat map should be grouped can be selected. When selecting a category in the dropdown all heat maps get updated immediately by sorting the rows by the selected category.

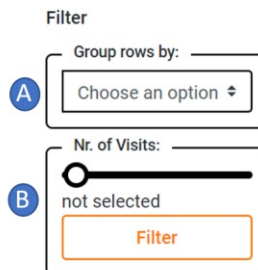


Figure 6.5: Filter View: (A) shows a dropdown input for the selection of a category by which the heat map should be grouped. (B) is a filter for the selection of patients that only attended a certain number of visits.

The second filter was additionally included and is a slider where the patients can be filtered by the number of visits they attended (see (B) in Figure 6.5). For example, some



patients dropped out of the study after the first visit and, therefore, had only two visits, the first visit and the last visit called “Study discontinuation” and when selecting two in the slider then only patients with two visits are shown which can be useful to analyze only patients that dropped out of the study.

### 6.3.4 Header

In the header (see Figure 6.6) of the prototype two additional features were added that were not defined in the mockups. Both of these features were suggestions from the domain expert. The first one is a search bar (A) that enables the user to search for certain questions in the two heat maps. Since both of the questionnaires consist of a large number of questions it is hard to distinguish between all of them in the heat map because of the size of the rectangles. Furthermore, sometimes users are only interested in certain questions about a certain topic, like pain or health and for such cases, it is helpful to have a search feature that makes it possible to search for “pain” for example and only questions including the word pain are then shown.

The second component of the header is a menu (B) that is connected to a new feature that was not defined in the mockups. With the menu, it is possible to switch between certain modes of the prototype which is explained in detail in the next section.

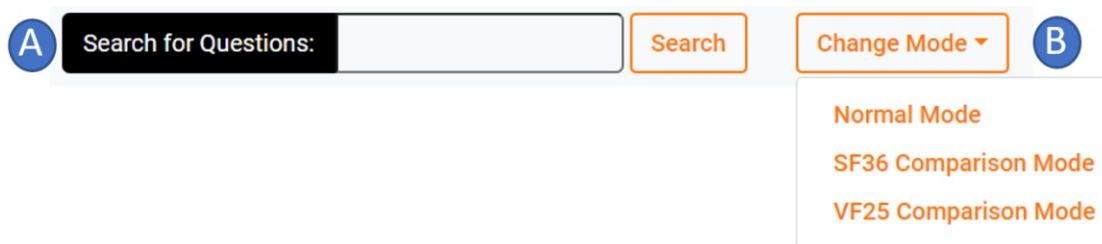


Figure 6.6: Header: (A) Search bar for the ability to search for certain questions and (B) is a menu for switching between the different modes of the prototype.

### 6.3.5 Comparison Mode

During the meetings with the domain expert, it became apparent that the comparison of groups of patients is an important task. With the initial version, the comparison between certain groups was only partly possible with the new dropdown input on top of the parallel coordinate charts. But as already criticized only a comparison between the two questionnaires is possible.

Therefore, the domain expert suggested adding the functionality to display one of the two questionnaires in both heat maps and parallel coordinate charts. This enables the user to filter for different groups of patients in both of the heat maps and parallel coordinate charts, making comparison of different patient groups within the same questionnaire possible.

## 6. PROTOTYPE IMPLEMENTATION

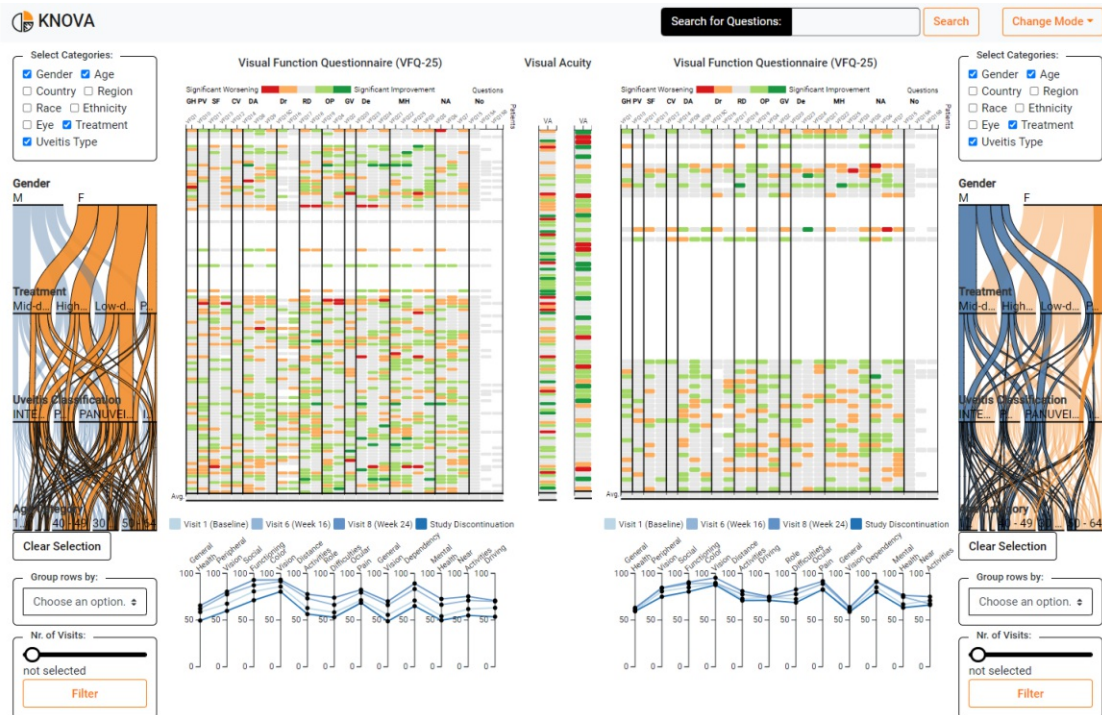


Figure 6.7: Comparison Mode: Additional mode that enables comparison of the same questionnaire. The screen is split into two halves where each chart visualizes the data from one questionnaire (the VF25 questionnaire in this case) and each half has its own filters which makes a comparison easier. This example visualizes all female patients on the left and all male patients on the right.

To provide this functionality different modes were implemented. The first one is called “Normal Mode” which is the mode that was already seen in the figures before and the two other modes are the comparison modes for the two questionnaires called “SF36 Comparison Mode” and “VF25 Comparison Mode”. Switching between those modes is possible through the menu in the header as explained in the previous section.

In Figure 6.7 the “VFQ Comparison Mode” can be seen. The structure of the two comparison modes is the same, the only difference is the data that is visualized in the heat maps and parallel coordinate charts.

The basic idea behind this mode is that in all charts the data from one questionnaire is shown (in Figure 6.7 the data from the VF25 questionnaire) and then comparisons can be done by applying filters for each half individually. The features and interaction techniques are all the same as in the normal mode. The only difference is that the screen is split in half and each half has its own parallel set filter, grouping filter and number of visits filter. An example can be seen in Figure 6.7 where in the left half the female patients were selected in the parallel set and on the right the male patients were selected

which makes a direct comparison of the data from males and females in the heat map as well as the parallel coordinate charts possible.

As can be seen in Figure 6.7 the visual acuity heat maps are duplicated as well and they are also affected by the filters from their half. Sorting is possible too in this mode but in the comparison mode, the sorting of one heat map does not affect all the others instead only the connected heat map is sorted. For example, if the left visual acuity heat map gets sorted then only the left questionnaire data heat map is sorted and on the right the same.

These were the main views and features of the final prototype. After the prototype was finished the evaluations were conducted with the process and results being discussed in the next chapter.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Evaluation

In this chapter, the final step of the thesis is presented which is the evaluation of the prototype. The objectives of the evaluation are:

1. to receive qualitative feedback from experts
2. to see how experts interact with the prototype to identify possible difficulties
3. to check if the visual encodings are clear
4. to check if the interaction techniques are clear

This chapter starts with the presentation of the preparation phase as well as the process of the evaluations themselves, followed by the results and feedback we received from the experts. In the end, the most important suggestions for improvements are summarized to provide an outlook on what could be improved in the future.

## 7.1 Preparation

As already mentioned in Chapter 2 Methodology, the evaluations were planned as task-oriented qualitative user studies combined with a thinking aloud approach with about 3-6 visualization experts [WAM<sup>+</sup>19; IIC<sup>+</sup>13]. In the end, we managed to conduct evaluations with five visualization experts.

The preparation of the evaluations started with a one-hour-long workshop where the domain expert, as well as three visualization experts, participated. The goal of this workshop was to define the tasks for the evaluation together with the domain expert since we wanted to know what type of tasks are the most important ones that should be included in the evaluation to see if the VA approach supports solving them. For the

tasks, we decided to use the classification proposed in the paper from Battle and Heer [BH19]. They classified tasks in four categories ranging from T1 to T4 starting from more general tasks to more open-ended ones. T1 are tasks intended to assess the data quality, whereas T2 and T3 intend to evaluate patterns and find relationships between parameters, and T4 tasks are more open-ended intended to analyze causality [BH19].

Prior to the workshop, we created a list of around 30 example tasks where at least one example task was present for each of the task classifications from Battle and Heer [BH19]. During the workshop, those tasks were discussed with the domain expert, and feedback was provided on which of the tasks are the most suitable ones for the evaluation as well as suggestions for some changes in the wording or phrasing. In the end, we managed to find eleven tasks which we used for the evaluation presented in the next section.

## 7.2 Evaluation Process

As already mentioned, the prototype was evaluated by five visualization experts. All participants had no prior knowledge about the dataset or the prototype. The evaluations were all conducted online via an online conferencing tool and each one lasted for around one hour. The advantage of using an online conferencing tool is that the evaluation can be easily recorded and via screen-sharing, all interactions with the prototype can be tracked and analyzed in detail afterward. Furthermore, during the whole evaluation, notes were taken in case if something went wrong with the recording.

Each evaluation was following the same structure:

1. Onboarding
2. Task Completion
3. Debriefing

### Onboarding

The onboarding started by introducing the participant into the prototype and by providing information about the dataset that was used for the prototype to give the participant a clue about the context of the thesis. After the introduction, three example tasks were presented as well as a way of solving those with the help of the prototype to make sure that the process of the evaluation is clear to them. The three tasks were the following ones which are three of the eleven tasks that were defined together with the domain expert:

1. What type of Uveitis is the least commonly enrolled from Fantasy Land?
2. How many females with the mid-dose show general improvement on question VFQ2? And how many show a general deterioration?

3. Compare the two treatment plans (placebo and high dose). Which one has the highest variation between the worst and best score for General Health? What is approximately the interval for each?

As already mentioned in the previous chapter we renamed the countries and regions to Disneyland names, like “Fantasy Land” or “Mickey’s Toon Town” to avoid any biases.

At the end of the onboarding, the participant was asked to open the prototype in the browser and to share the screen. After opening the prototype and sharing the screen the participant got some time to try the prototype out on their own to familiarize themselves with the prototype and all its components and interaction techniques. When planning the evaluation we first did not consider this step of the participant getting familiar with the prototype but after the first evaluation we received the feedback that solving the tasks immediately after the onboarding without having time to try out the prototype was overwhelming. Therefore, we added this step in the onboarding of all the remaining four evaluations.

The final step was then to ask the participant to think aloud during the whole evaluation if possible to get even more insights by hearing what the person thinks while using the prototype. Since the prototype is very feature-rich the onboarding took around 20 minutes.

### Task Completion

After the onboarding was finished and everything was set up the actual evaluation started. As already mentioned, three of the eleven predefined tasks were used for the onboarding, and the remaining eight were used for the evaluation itself. Therefore, we had a list of eight tasks (see Table 7.1) which we revealed one at a time to not cause any distractions. Furthermore, the questions were ordered by complexity as well as the classification from Battle and Heer [BH19], starting with T1 and T2 questions, and closing with a T3 task. The reason for this order was to ensure a learning curve where the users, in the end, are reaching a level of familiarity that allows them to solve even more complex tasks.

### Debriefing

The last step of the evaluation, after all tasks were finished, was to ask the participant about their general impression of the prototype as well as if they have any kind of feedback or improvement suggestions regarding the prototype or the evaluation itself.

## 7.3 Results

The five evaluations were conducted in the range of about two weeks and after all of them were finished the analysis started.

First, all recordings were rewatched to make sure that all important comments, improvement suggestions, as well as feedback, were captured. Furthermore, the task completion



Task	Classification	Task Description
1	T1	What is the percentage of young people (19-29) having Panuveitis?
2	T1	What (imaginary) country is the most represented in the study?
3	T1	Which treatment group had the least variation in the SF-36 category 'Pain' (SFQ7, SFQ8)?
4	T2	Compare the two treatment groups (placebo and high dose) according to Visual Acuity. Which group has the largest proportion of patients who did not show either improvement or deterioration?
5	T1	What type of Uveitis is the most commonly found in patients from Mickey's Toon Town?
6	T2	Which treatment group of participants had the highest number of dropouts with 2 visits (Visit 1 + Study Discontinuation)?
7	T2	Is it possible to correlate a decrease of the pain health parameters to a decrease of the visual functioning parameters?
8	T3	What are the parameters that are most commonly associated to a study discontinuation?

Table 7.1: List of tasks that the users were asked to solve including the classifications from Battle and Heer [BH19].

time for each task was noted. The second step was the extraction and summary of the most important observations as well as improvement suggestions which are presented in the next sections.

### 7.3.1 Completion Times

When analyzing the completion times depicted for each task and participant in Table 7.2 some observations can be made. For example, the order of the tasks by their classification which ranges from easier tasks to more complex ones is reflected in the timings. The average timings are increasing and especially in the end the timings are getting longer which makes sense since the tasks 6 to 8 are more complex as well as more open-ended and allow for interpretations resulting in a longer completion time.

Another aspect that is reflected in the timings is that the first participant (P1) needed more time for some tasks compared to the others, for example for the tasks 1 and 3. This reflects the change that was made after the first evaluation because the first participant did not get a few minutes to familiarize themselves with the prototype on their own and started solving the tasks immediately after the introduction. Therefore, the first learning curve was missing for the first participant and the completion times seem to support the

Task/Participant	P1	P2	P3	P4	P5	Average (in minutes)
Task 1	04:22	01:00	00:36	00:38	00:44	<b>01:28</b>
Task 2	01:31	01:30	00:26	00:40	00:30	<b>00:55</b>
Task 3	12:00	05:50	01:30	02:05	01:32	<b>04:35</b>
Task 4	03:40	02:31	02:41	00:45	02:45	<b>02:28</b>
Task 5	03:00	00:40	01:00	01:30	00:35	<b>01:21</b>
Task 6	05:00	06:27	04:00	02:35	02:58	<b>04:12</b>
Task 7	06:00	03:44	03:32	03:55	04:40	<b>04:22</b>
Task 8	01:46	07:00	09:00	05:10	12:45	<b>07:08</b>

Table 7.2: Task completion time per task and participant as well as the average time per task. The participants are abbreviated with P1 to P5.

decision of changing the onboarding and make clear how important such a small change can be.

Something that stands out is that Task 3 has a rather high average completion time compared to the previous tasks as well as the two following tasks which does not fit the mentioned pattern of the increasing completion time with increasing task number. The reason for the higher average value are the timings from the first two participants. They both had difficulties with this task. For example, the first participant was confused about where to find the 'Pain' category because the subscore labels on top of the chart are just abbreviated and in some cases, the spacing between the different subscores is very low which makes the distinction between the subscores difficult. Furthermore, P1 was first not completely sure how to solve this task properly and started to try all the different features of the prototype which was then a short phase of trying the prototype out as the others had in the beginning which is also a reason for the higher completion time. When observing the second participant solving the third task the main reason for the higher completion time was that the comparison mode was not clear at first but after a short recap of the comparison mode the person solved the task rather quickly.

### 7.3.2 Observations

During the onboarding some of the participants already asked questions to make sure that they understood everything correctly, for example what the white rows in the heat map mean or asked for more details regarding the parallel set, if it is possible to select multiple strands, etc. Nevertheless, when observing the first minutes the participants had to familiarize themselves with the prototype on their own it seemed that the structure as well as the general interactions were understood rather quickly. Furthermore, the visual encodings seemed to be clear which was one important objective of our evaluation.

Other general observations we made were that the tooltip of the heat map was in the way instead of being a help most of the time. Especially when selecting rows for the counting feature it was annoying and made the selection more difficult. Moreover, for none of

the tasks the tooltip was necessary to use but it was still very present. One participant suggested for example to provide the user the possibility of turning the tooltips off.

The most confusing part of the evaluations was Task 6 in combination with the number of visits filter. The main problem with this task and the filter was that the participants were lacking prior domain knowledge. Therefore, they had their difficulties with understanding how to figure out if a patient dropped out and especially what the “Study Discontinuation” visit means. It was not clear if “Study Discontinuation” is the same as a study dropout or if there is a difference. Furthermore, the number of visits filter was confusing as well since it is a slider input which was seen as a range selector, meaning that when selecting, for example, three visits it was thought that patients with 0, 1, 2 and 3 visits are selected and not just the patients with three visits as it actually is.

When observing the interaction with the parallel set as well as the thoughts of the users some issues arose. For example, for some users it was not clear that the parallel set always stays the same and is not affected by the filters. Furthermore, one participant suggested to make the highlighting of the selected strands in the parallel set more present, for example, by making the font of the selected category bold and increase the transparency of the strands that are not selected. Another problem recognized in the parallel set is that if a category on the bottom of the parallel set is selected that has no following strands then no highlighting at all is provided and, therefore, some were not sure if the selection was actually working or not. In general, the interaction with the parallel set seemed to be sometimes a bit hard due to the small size because the selection of small strands was difficult and the labels of the categories are sometimes not shown or just abbreviated which makes it even harder. One participant was first confused where to find the treatment groups in the parallel set since the word “dose” was not shown at all or another confusing aspect was that the age categories were not ordered in the parallel set.

However, since this sizing problem was very present during the first two evaluations we decided to implement the ability of expanding the patient selection view which was explained in Section 6.3.1 and the last three participants were able to use this feature. Therefore, we had the ability to compare whether this change was helpful or not and it turned out that the new feature was used very often during the evaluations of the last three users and proved to be effective.

Regarding the parallel coordinates chart it was observed that the sequential color scheme makes it difficult to distinguish between the different lines.

Moreover, two minor bugs were revealed during the evaluations. The first one affects the reordering feature of the parallel set visualization. When reordering the categories and dragging a category too far to the top parts of it are not visible anymore. The second bug is that when selecting rows for the counting functionality white rows that indicate that the questionnaire was not answered by this patient cannot be selected which sometimes results in wrong counting results.

Regarding the tasks and their wording we could identify some difficulties and confusions as well. For example, some of the participants were not sure what “visual functioning

parameters” mean in Task 7. For the last task some participants were not sure what type of parameters they were supposed to find.

### 7.3.3 General Feedback

The general feedback we received from the participants was very positive. Although some of them commented that they first felt a bit overwhelmed, in the end, they got used to the prototype very quickly and enjoyed playing around with it. Furthermore, they noted that the visualization is very feature-rich and provides a lot of visualizations but still looks tidy and has well-working interactions.

Some general improvement suggestions were, for example, to make the font size a bit larger since it was sometimes a bit hard to read. Regarding the grouping feature, it was suggested to maybe add the possibility of grouping by visit number or by worsening/improvement. Furthermore, one participant commented on the subscore labels on top of the heat map to add more space between the labels or to move them to the bottom of the heat map because some did not see them at first or were not sure what those labels mean.

Another participant suggested adding the ability to zoom into the heat maps because it was sometimes a bit hard to analyze the heat map due to the huge amount of data and a zooming feature could help to focus more on smaller parts of the data. Furthermore, it was suggested to provide the possibility of being able to perform the counting feature multiple times at once without removing the previous count.

We also received some feedback regarding the evaluation itself. As already mentioned, the first participant gave the feedback that a short time for trying the prototype out on their own was missing which was the reason why we added this to the following evaluations. Furthermore, it was noted that many of the tasks involved mainly the parallel set visualization, and only one task was regarding the comparison mode. This means that the tasks could have been more balanced in focusing on all the aspects of the prototype evenly.

In the end, we also asked them if they had the impression that some of the difficulties they had were caused by the fact that they had no prior knowledge about the dataset and its domain. Some of them agreed with this hypothesis pointing out that especially the part with the visits would have been a lot clearer with domain knowledge.

To sum it up, the evaluation was successful in gaining new insights regarding the prototype and its acceptance of the users. Furthermore, when considering our objectives from the beginning it can be said that they were mostly fulfilled. In this section, the results from the first objective “receive qualitative feedback from experts” were presented. The second one, analyzing how the experts interacted with the prototype as well as identifying possible difficulties was fulfilled as well since we were able to observe the participants when interacting with the prototype and some difficulties were identified, as the problem of the very prominent tooltip, for example. The third objective regarding the visual encodings was fulfilled as well since the visual encodings were all clear and no difficulties

were identified that were caused due to the encoding. The last objective was regarding the interaction techniques and those were also mostly clear but some improvement is still possible. For example, the counting feature could be simplified by making it a brushing interaction instead of a two-click interaction.

### 7.3.4 Identified Improvements

To summarize the results of the evaluation process a list of possible improvements for the future is provided:

- Increase the font size to ensure better readability.
- Change color coding of lines in parallel coordinate charts to be able to distinguish them better.
- Enable the possibility of selecting that only certain visits should be displayed in the parallel coordinate charts, for example by clicking on the color rectangles in the legend of the chart.
- Make the highlighting of the selected strands in the parallel set more prominent by using a bold label as well as increasing transparency of the other strands.
- Add the possibility of deactivating heat map tooltips.
- Provide the possibility of being able to zoom into the heat map.
- Extend the counting feature to allow the selection of multiple countings at once.
- Improve number of visits filter, by either changing it to a number input instead of a slider or a different solution.
- Include functionality that makes the selection of patients that dropped out possible.

In the end, we implemented one of the suggestions listed above which is the number of visits filter since this was the feature that caused the most confusion and was important for many tasks. The first consideration was if it is enough to change the slider input to a number input to avoid confusion whether the selected number in the slider input is a fixed number or if it indicates a range. The second consideration was to replace the slider input with a dropdown input where the user can select whether to show only patients that dropped out of the study or only patients that completed the study. The advantage of the second idea is that it solves the issue that it was not clear what it means if someone filters for two visits or three visits. When discussing those considerations with the domain expert we came to the conclusion to implement the second idea since this one is more clear and straight-forward.

The result of this change can be seen in Figure 7.1 where on the left the old number of visits filter is shown and on the right the new version. The new version consists of a

dropdown input with three options. The first one is “All Patients” which is the option that is selected by default showing all patients. The others are “Discontinued Patients” filtering for all patients that dropped out of the study without considering after which visit they dropped out and the last one is “Completed Patients” which filters only for patients that completed the whole study.



Figure 7.1: Improvement of the number of visits filter: Old version using a slider input on the left and on the right the new version using a dropdown input.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.



# Conclusion & Future Work

## 8.1 Summary

The goal of this thesis was to provide an overview of the current state of the art in the KAVA research field with a focus on approaches dealing with health care data. Since the number of approaches in this area is rather low we tried to fill this gap by conducting problem-driven visualization research as well as designing and developing a new approach that visualizes data from a clinical trial where the efficacy of a medication for the treatment of the eye disease Uveitis was analyzed. In this approach, general patient data were visualized as well as data from questionnaires that assessed the quality of life of the patients. The domain knowledge included in our approach was declarative, provided by the questionnaires. The declarative knowledge is, on the one hand, the questionnaire itself with its structure and questions, and, on the other hand, the predefined subscores that reflect certain aspects of the patients' quality of life.

In the end, a qualitative task-oriented user study was conducted with five visualization experts to get feedback as well as to evaluate whether the encodings and interaction techniques are clear or if some parts need to be improved further. Furthermore, during the whole design and development phase, a domain expert provided us with constant feedback regarding the approach as well as providing detailed knowledge about the dataset which we integrated and considered.

The thesis was based on the following research questions and subquestions which we were able to answer after finishing the whole process starting from researching already existing approaches to developing, and evaluating an own approach.

**RQ** *How can an interactive exploration environment supported by the integration of domain knowledge help and assist medical researchers to explore health care data of whole patient cohorts?*

To answer this question we developed a new approach that visualizes data from a clinical trial and evaluated it afterward. As the evaluations have shown as well as the feedback we received from the domain expert a VA approach has great potential for the analysis and exploration of health care data. Especially in the field of clinical trials the resulting data are often in a tabular form and only analyzed statistically to check whether the endpoints of the trial were met to be able to decide on the approval of the medication [Lam20].

Nevertheless, a more detailed analysis could yield new insights besides the results of the efficacy evaluation and for such analysis, VA approaches can be a great help. They can help to get another view on the data as well as to explore it in more detail. Furthermore, by the integration of domain knowledge, the VA approach is enhanced even more which makes such approaches a real enrichment for health care data analysis.

**SQ1** *Which insights into a dataset and its structure can be gained when using a tool, like Tableau to explore a dataset visually?*

Tableau makes it possible to create fast and easy visualizations of a dataset which is especially helpful at the beginning of a project where the dataset, as well as its structure, is not known. In our thesis, the data exploration phase using Tableau was very valuable since we could gain a great understanding of the dataset. Furthermore, it helped to gain insights, like which encodings fit best to the dataset by trying out different types of visualizations. For a dataset like ours that includes many different parameters, it is also helpful to identify which parameters of the data are more important than others to decide on which ones to focus on. While exploring the data in Tableau we also recognized that the VFQ25 questionnaire data were missing for a large number of patients which is something that can be identified very easily with a visual representation of the data.

**SQ2** *Which visualization techniques are most suitable for the visualization of health care data?*

After researching for existing approaches as well as exploring the dataset using Tableau we could identify several visualizations that are suitable for our dataset. Since our dataset deals with data from a patient cohort with a large number of patients, it was important to have a visualization that can show aggregated data in a compact form which we reached by using the heat maps. Furthermore, health care data are often time-oriented and, therefore, visualizations are needed that can visualize the timely evolution of the data which we reached by using the parallel coordinate charts. Finally, the parallel set visualization was used for the visualization of categorical data, like patient-related data.

**SQ3** *Which interaction techniques are necessary and suitable for the visualization of health care data?*

As already mentioned, health care data includes often a large number of parameters as well as data from many patients. Therefore, interaction techniques are very important for visualizations of health care data since they make it possible to narrow the data down, for example, by using filters to be able to analyze certain aspects in more detail.

Therefore, we include interaction techniques, like different filters as well as tooltips that help to get more details on demand. Furthermore, the counting feature in our approach proved to be very helpful when observing the participants during their evaluation.

## 8.2 Future Work

Considering the feedback received from the domain expert as well as the results of the evaluations which were very positive the developed approach seemed to be a good starting point for further research in this direction. The next steps in the future could be to implement the improvement suggestions that came up during the evaluations. Furthermore, evaluations with domain experts would be important to get their opinion regarding the approach as well as to analyze whether their impressions, feedback, and improvement suggestions are similar to the ones from the visualization experts or if they differ.

Since this approach dealt with a certain dataset it could be interesting to analyze whether it also works with other datasets.

Another interesting aspect would be to enhance the approach by integrating more domain knowledge, especially in the form of procedural knowledge since the current approach only deals with declarative knowledge.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# List of Figures

2.1	Nested Model [Mun09]: Base approach for the design process. . . . .	6
2.2	Design Triangle [MA14]: Components that we consider during our design process. . . . .	6
2.3	Snippet from the dataset that will be used as a base for the approach. . .	7
2.4	Schematic timeline of the thesis: All main milestones (blue color) are arranged in their temporal order as well as their main sub-tasks (grey color) that are conducted in between. . . . .	8
3.1	CareVis [AM06]: Visualization of treatment plans combined with patient data from a single patient. On the left, the structure of the treatment plan is visualized. Next to it, the executed actions of the plan are visualized as well as certain parameters the plan affected and on top is the QuickView panel providing information regarding the most important patient parameters. . .	11
3.2	CareCruiser [GAK <sup>+</sup> 11]: Visualization of treatment plans and multiple patient records. On the left, the treatment plan is visualized and on the right, the patient parameters are presented as well as the executed actions using diamonds. . . . .	12
3.3	Gnaeus [FUA <sup>+</sup> 15]: Visualization of the treatment plan structure and patient data of whole cohorts. (A) shows the hierarchical structure of the treatment plan whereas (B) displays the structure of the selected subplan. (C) is a streaming box-plot visualization of the aggregated parameter data for the selected subplan and (D) a pixel-based visualization of the relevant patient parameter for each patient. (E) depicts the visualization of the aggregated data regarding the executed actions and (F) the grouping lens interaction technique that visualizes the distribution as stacked bars. (G) shows the vertical fish-eye distortion available for detailed patient analysis. . . . .	13
3.4	Web-based visualization [BPG <sup>+</sup> 15]: Sankey diagram of subprocesses that visualize the flows between executed care activities. . . . .	14
3.5	Web-based visualization [BPG <sup>+</sup> 15]: Conformance flowchart visualizing all possible paths of the guideline. (Figure modified) . . . . .	15
3.6	Lifelines2 [WPS <sup>+</sup> 09]: An approach that visualizes event sequences of multiple patients. (a) shows the visualization of event sequences per patient and (b) a temporal summary. (c) are controls for alignment, ranking, and filtering, and (d) controls for the temporal summary and navigating groups. . . . .	16
		83

3.7	LifeFlow [WGP <sup>+</sup> 11]: Visualization of event sequences. On the left, controls for aligning, filtering, and ranking are present as well as a legend for selecting certain event types. In the center, a tree-based visualization represents the event sequences and on the right, LifeLines2 is used for the visualization of event sequences of single patients. . . . .	17
3.8	Timeline-based visualization [DJC17]: Timeline-based visualization of a patient history. . . . .	18
4.1	A world map visualizing the distribution of the various Uveitis classifications among the different countries. . . . .	22
4.2	Bar charts visualizing distributions by Uveitis classification: (A) shows the distribution among the age categories and (B) shows the distribution among female and male patients. . . . .	23
4.3	“Packed bubbles” visualizing groups of patients based on the Uveitis classification, etiology, and gender. . . . .	24
4.4	Heat map for the SF36 questionnaire answers of all patients per visit. . .	25
4.5	Questionnaire data encodings: (A) Sparkline chart, (B) Color-coded lines. . .	26
4.6	Heat map visualization of the aggregated data for both questionnaires. . .	27
4.7	Bar chart visualizing the subscores from the SF36 questionnaire for both male and female patients. . . . .	28
4.8	Radar chart visualizing the subscores from the SF36 questionnaire per visit. . .	29
4.9	Dashboard visualizing general patient data: (A) world map visualizing the Uveitis classification per country, (B)-(D) bar charts visualizing the Uveitis distribution per gender, age category, and etiology, (E) filter pane, (F) table displaying patient distribution per treatment group. . . . .	31
4.10	Dashboard visualizing questionnaire results: (A) heat map visualizing the change from the first visit to the last visit, (B) heat map visualizing the answers of all visits, (C) filter and legend pane. . . . .	32
4.11	Dashboard visualizing subscore data from SF36 questionnaire: (A) bar chart visualizing summary scores per gender, (B) summary scores per Uveitis classification, (C) bar chart for subscores per gender, (D) radar chart visualizing subscores for each visit, (E) bar chart visualizing subscores per Uveitis classification, (F) filter and legend pane. . . . .	33
4.12	Dashboard visualizing subscore data from VF25 questionnaire: (A) Radar chart visualizing subscores per visit, (B) bar chart visualizing subscores per gender, (C) another bar chart that visualizes the subscores per Uveitis classification, (D) filter and legend pane. . . . .	34

5.1	First mockup of own approach: The view is vertically split where (A) includes visualizations of the SF36 questionnaire data and (B) visualizations of the VF25 questionnaire data. Both halves include the same visualizations, firstly, the radar chart (C) visualizing how the average subscores have changed from visit to visit, and, secondly the heat map (D) visualizing how the answers from each patient for each question have changed over time. (E) is a filter pane that includes multiple filters that affect both, the radar chart as well as the heat map. (F) illustrates the lens interaction technique that reveals the data behind the aggregated values which are the answers of the patients for the questions at all visits. . . . .	45
5.2	Updated mockup: (A) The rows in the heat map are separated using lines to indicate which questions belong to which subscore. (B) A dropdown input that enables sorting of the heat map by certain characteristics, like gender, Uveitis classification, etc. . . . .	47
5.3	Updated mockup with details on demand: (A) shows the tooltip that is displayed when hovering over a rectangle containing patient-related data, like age, gender, country, etc. as well as a line chart visualizing how the patient answered the selected question at each visit. In the radar chart certain subscores were selected which are highlighted in the heat map by greying the others out. . . . .	48
5.4	Updated mockup with a new feature: (A) shows sparklines that visualize how each patient answered each question at the different visit which is only displayed if enough space is given through filtering. . . . .	49
5.5	Updated mockup: All subscore labels were added to the questions (A), (B) shows a tooltip with the full question text that is displayed on hovering over the “i” icon. (C) Axis indicators were added to the heat map as well as the “[shuffled]” indicator and (D) is a new column at the end of the heat map visualizing the average value over all patients for each question. . . . .	51
5.6	New approach - “Cluster-based” selection: (A) Parallel set visualizing how the patients are distributed among categories, like gender, region, or Uveitis classification which can be used as a filter for the other charts by clicking on one of the boxes. (B) is the currently collapsed “filter-based” selection and (C) provides some general statistics about the data. (D) is the main view that is split in half for the two questionnaires where each of them consists of a heat map and a parallel coordinates chart. (E) displays a tooltip providing patient-related data as well as data about the questionnaire and (F) is a filter pane with some filters as well as the sorting feature. . . . .	53
		85



5.7	New approach - “Filter-based” selection: (A) indicates the currently collapsed “cluster-based” selection. (B) is the “filter-based” selection with some dropdown inputs and checkboxes that are responsible for which patients are displayed in the other charts. (C) is again the area with some general statistics about the dataset. The main part (D) is the same as in the other part with the heat map and parallel coordinates chart for each questionnaire as well as the tooltip (E). With the dropdown input (F) it can be selected how the patients should be sorted in the heat map which is in this case by gender which is also displayed in the heat map. . . . .	54
6.1	Prototype layout: The user interface is horizontally split into three views: (A) the patient selection view, (B) the main view, and (C) the filter view. The header (D) includes a search bar and a menu for changing the mode of the prototype. . . . .	59
6.2	Patient selection view: (A) Multiple checkboxes that define which categories are shown in the parallel set (B). (C) is a tooltip showing how many patients belong to the selected set and with the “Clear Selection” button (D) the selection is reset. With the arrow on the top right (E) the patient selection view can be expanded or condensed. In this figure the view is expanded. .	60
6.3	Main View - Heat maps: (A) displays the newly added heat map for the visual acuity scores using the same encoding as the other heat maps. (B) is a tooltip containing patient-related information as well as a line chart visualizing the concrete answers per visit. (C) shows that the visual acuity heat map was sorted and since the grouping is active the sorting is done for each group separately. In (D) the counting feature is depicted. . . . .	62
6.4	Main View - Parallel coordinate charts: (A) shows the tooltip that is displayed on hovering over one of the dots providing information about the visit type as well as the actual value and the number of patients that were used to calculate the average value. (B) shows dropdown inputs that are shown if the grouping filter is active. It enables the user to switch between certain groups of patients, for example, patients from the placebo group and the low-dose treatment group. . . . .	63
6.5	Filter View: (A) shows a dropdown input for the selection of a category by which the heat map should be grouped. (B) is a filter for the selection of patients that only attended a certain number of visits. . . . .	64
6.6	Header: (A) Search bar for the ability to search for certain questions and (B) is a menu for switching between the different modes of the prototype. . .	65
6.7	Comparison Mode: Additional mode that enables comparison of the same questionnaire. The screen is split into two halves where each chart visualizes the data from one questionnaire (the VF25 questionnaire in this case) and each half has its own filters which makes a comparison easier. This example visualizes all female patients on the left and all male patients on the right.	66
86		

7.1 Improvement of the number of visits filter: Old version using a slider input on the left and on the right the new version using a dropdown input. . . . 77



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# List of Tables

3.1	Tabular summary of all advantages and disadvantages of each presented approach. . . . .	20
6.1	Mapping scheme of the visual acuity difference value to the color scale of the heat map ranging from red to green. . . . .	61
7.1	List of tasks that the users were asked to solve including the classifications from Battle and Heer [BH19]. . . . .	72
7.2	Task completion time per task and participant as well as the average time per task. The participants are abbreviated with P1 to P5. . . . .	73



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar  
The approved original version of this thesis is available in print at TU Wien Bibliothek.

# Bibliography

- [AM06] Wolfgang Aigner and Silvia Miksch. Carevis: integrated visualization of computerized protocols and temporal patient data. *Artificial Intelligence in Medicine*, 37(3):203–218, 2006.
- [AMTB05] Wolfgang Aigner, Silvia Miksch, Bettina Thurnher, and Stefan Biffl. Planninglines: novel glyphs for representing temporal uncertainties and their evaluation. In *Ninth International Conference Information Visualisation (IV)*, pages 457–463, July 2005.
- [BH19] Leilani Battle and Jeffrey Heer. Characterizing exploratory visual analysis: a literature review and evaluation of analytic provenance in tableau. *Computer Graphics Forum*, 38:145–159, June 2019.
- [BKH05] Fabian Bendix, Robert Kosara, and Helwig Hauser. Parallel sets: visual analysis of categorical data. In *IEEE Symposium on Information Visualization*, pages 133–140, 2005.
- [BPG<sup>+</sup>15] Rahul C. Basole, Hyunwoo Park, Mayank Gupta, Mark L. Braunstein, Duen Horng Chau, and Michael Thompson. A visual analytics approach to understanding care process variation and conformance. In *Proceedings of the 2015 Workshop on Visual Analytics in Healthcare*, 6:1–6:8. Association for Computing Machinery, 2015.
- [CG15] Jesus Caban and David Gotz. Visual analytics in healthcare—opportunities and research challenges. *Journal of the American Medical Informatics Association: JAMIA*, 22:260–262, 2015.
- [Col02] Anne L Coleman. Development of the 25-item national eye institute visual function questionnaire. *Evidence-Based Ophthalmology*, 3(1):58–59, 2002.
- [DJC17] Filip Dabek, Elizabeth Jimenez, and Jesus J. Caban. A timeline-based framework for aggregating and summarizing electronic health records. In *IEEE Workshop on Visual Analytics in Healthcare (VAHC)*, pages 55–61, 2017.
- [FHR<sup>+</sup>14] Paolo Federico, Stephan Hoffmann, Alexander Rind, Wolfgang Aigner, and Silvia Miksch. Qualizon graphs: space-efficient time-series visualization with qualitative abstractions. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pages 273–280, 2014.

- [Fra92] I Conceptual Framework. The mos 36-item short-form health survey (sf-36). *Med Care*, 30(6):473–83, 1992.
- [FUA<sup>+</sup>15] Paolo Federico, Jürgen Unger, Albert Amor-Amorós, Lucia Sacchi, Denis Klimov, and Silvia Miksch. Gnaeus: utilizing clinical guidelines for a knowledge-assisted visualisation of ehr cohorts. In *Proceedings of the Euro-Vis Workshop on Visual Analytics (EuroVA)*, pages 79–83. Eurographics, 2015.
- [FVO<sup>+</sup>17] Claudia Fabiani, Antonio Vitale, Ida Orlando, Marco Capozzoli, Fiorella Fusco, Francesco Rana, Rossella Franceschini, Jurgen Sota, Bruno Frediani, Mauro Galeazzi, et al. Impact of uveitis on quality of life: a prospective study from a tertiary referral rheumatology-ophthalmology collaborative uveitis center in italy. *The Israel Medical Association journal: IMAJ*, 19(8):478–483, 2017.
- [FWR<sup>+</sup>17] Paolo Federico, Markus Wagner, Alexander Rind, Albert Amor-Amorós, Silvia Miksch, and Wolfgang Aigner. The role of explicit knowledge: a conceptual model of knowledge-assisted visual analytics. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 92–103, 2017.
- [GAK<sup>+</sup>11] Theresia Gschwandtner, Wolfgang Aigner, Katharina Kaiser, Silvia Miksch, and Andreas Seyfang. Carecruiser: exploring and visualizing plans, events, and effects interactively. In *Proceedings of the IEEE pacific visualization symposium (PacificVis)*, pages 43–50, 2011.
- [GES16] Zuguang Gu, Roland Eils, and Matthias Schlesner. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, 32(18):2847–2849, 2016.
- [GWFI14] Pascal Goffin, Wesley Willett, Jean-Daniel Fekete, and Petra Isenberg. Exploring the placement and design of word-scale visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 20:2291–2300, December 2014.
- [IIC<sup>+</sup>13] Tobias Isenberg, Petra Isenberg, Jian Chen, Michael Sedlmair, and Torsten Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, 2013.
- [KP15] Simone Kriglstein and Margit Pohl. Choosing the Right Sample? Experiences of Selecting Participants for Visualization Evaluation. In W. Aigner, P. Rosenthal, and C. Scheidegger, editors, *EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (EuroRV3)*. The Eurographics Association, 2015.
- [KWEH14] Noreen Kamal, Samuel Wiebe, Jordan Engbers, and Michael Hill. Big data and visual analytics in health and medicine: from pipe dream to reality. *Journal of Health and Medical Informatics*, 5:e125, February 2014.



- [Lam20] Jean-Baptiste Lamy. Comparison of four visual analytics techniques for the visualization of adverse drug event rates in clinical trials. In *24th International Conference Information Visualisation (IV)*, pages 344–349, 2020.
- [Lew82] Clayton Lewis. *Using the “thinking-aloud” method in cognitive interface design*. IBM TJ Watson Research Center Yorktown Heights, NY, 1982.
- [LSA<sup>+</sup>16] Roger A Leite, Lucas Mello Schnorr, Jurandy Almeida, Bruna Alberton, Leonor Patricia C Morellato, Ricardo da S Torres, and João LD Comba. Phenovis—a tool for visual phenological analysis of digital camera images using chronological percentage maps. *Information Sciences*:181–195, 2016.
- [MA14] Silvia Miksch and Wolfgang Aigner. A matter of time: applying a data–users–tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics*, 38:286–290, 2014.
- [Man] CM Mangione. The national eye institute 25-item visual function questionnaire (vfq-25).
- [Mun09] Tamara Munzner. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928, 2009.
- [PL20] Bernhard Preim and Kai Lawonn. A survey of visual analytics for public health. *Computer Graphics Forum*, 39(1):543–580, 2020.
- [Saa08] M Joan Saary. Radar plots: a useful way for presenting multivariate health care data. *Journal of clinical epidemiology*, 61(4):311–317, 2008.
- [SMM12] Michael Sedlmair, Miriah Meyer, and Tamara Munzner. Design study methodology: reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012.
- [TC06] James J Thomas and Kristin A Cook. A visual analytics agenda. *IEEE Computer Graphics and Applications*, 26(1):10–13, 2006.
- [WAM<sup>+</sup>19] Emily Wall, Meeshu Agnihotri, Laura Matzen, Kristin Divis, Michael Haass, Alex Endert, and John Stasko. A heuristic approach to value-driven evaluation of visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):491–500, 2019.
- [War01] John Ware. Sf-36 health survey update. *Spine*, 25:3130–9, January 2001.
- [WGP<sup>+</sup>11] Krist Wongsuphasawat, John Alexis Guerra Gómez, Catherine Plaisant, Taowei David Wang, Meirav Taieb-Maimon, and Ben Shneiderman. Lifeflow: visualizing an overview of event sequences. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1747–1756. ACM, 2011.
- [Wha13] William Whang. *Medical outcomes study*. In *Encyclopedia of Behavioral Medicine*. Marc D. Gellman and J. Rick Turner, editors. Springer New York, New York, NY, 2013, pages 1210–1211.

- [WMK93] John Ware, Kosinski MA, and S.D. Keller. Sf-36 physical and mental health summary scales: a user's manual. 8:23–28, January 1993.
- [WPQ<sup>+</sup>08] Taowei David Wang, Catherine Plaisant, Alexander J Quinn, Roman Stan-chak, Shawn Murphy, and Ben Shneiderman. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 457–466. ACM, 2008.
- [WPS<sup>+</sup>09] Taowei David Wang, Catherine Plaisant, Ben Shneiderman, Neil Spring, David Roseman, Greg Marchand, Vikramjit Mukherjee, and Mark Smith. Temporal summaries: supporting temporal categorical searching, aggregation and comparison. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1049–1056, 2009.