

Patient Reported Outcome Visualization (PROVis): Design Study to Support Pediatric Rheumatologists



Thijs Rutjes (5494656)

Utrecht University

Department of Information and Computing Sciences

Business Informatics

First supervisor: Dr. Evanthia Dimara

Second supervisor: Dr. Hanna Hauptmann

Daily supervisors: Martijn Doeleman & Dr. Sytze de Roock

Abstract

Patient Reported Outcomes (PRO) is any report of the status of a health condition that comes directly from the patient without interpretation by the clinician. PROs contain valuable information that a doctor cannot easily observe, such as symptoms experienced in day-to-day life. However, clinicians face a difficulty to use the format of PROs, typically gathered via a questionnaire, preferring to interact with patients directly. Consequently, clinicians appear repeating questions already gathered in the PRO data, while patients experience their detailed PRO reporting as a time-wasting activity. Existing attempts at improving the format of PROs have not focused on creating a solution that meets the time requirements of clinicians. This study aims to learn how to improve the format of PROs by means of visualization, so PROs can be effectively integrated in clinician workflows. To achieve such integration, this study followed the nine-stage “design study” methodology, where visualization researchers directly investigate a real-world problem in a target domain, design a validated visualization solution for it, and reflect on lessons learned to refine guidelines. This study was thus conducted at a pediatric rheumatology department of a Dutch hospital. Fly-on-the-wall observation of a pediatric rheumatologist followed by a semi-structured interview indicated that clinicians need a quick overview of the progression of PRO results of an individual patient over time while comparing these results to medication changes. Interviews of four patients indicated that patients need their doctors during their meeting to show that they understood and remember their PRO data well to the extent that they pose relevant rather than repetitive questions. Based on those observations, PROVis was developed, a visualization system that supports the clinicians in quickly learning the most important PRO data in a narrative summary and allowing for identified attributes to be plotted over time through line graphs. The visualization system is centered around hospital visit dates, clearly illustrating whether the patient filled in their questionnaire for each visit. PROVis has been evaluated with four pediatric rheumatologists and a nurse. A thematic analysis of the evaluation has shown that PROVis allows clinicians to get a quick grasp of the PRO results as well as an understanding of the chronology of the PRO data. Future renditions of PROVis should include comparisons of PRO data with clinician observations. The results of this study stress the importance of visualizations for clinicians taking the limited time available during clinics into account. Visualizations of PRO data should further illustrate the results over time, while also allowing comparisons with the changes in medication and the measurements of the clinicians.

1 Introduction

1.1 Motivation

Clinicians observe their patients to find out how severe their disease is and make treatment decisions based on that. However, studies show a significant difference between the observations from clinicians and how the patients describe the severity of their symptoms [1]. For example, a clinician may think a patient is doing well because no joint inflammation could be observed with objective measures. Meanwhile, the patient could still experience pain in their joints in their day-to-day life, contradicting the observations of the clinician. To tackle this discrepancy, there has been a shift to more patient-centric healthcare with the Patient Reported Outcome [2].

A Patient Reported Outcome (PRO) is any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else [3]. PROs provide unique information on the impact of a medical condition and its treatment from the patients' perspective [4]. Going back to the earlier example of painful joints, a patient could report their experienced pain severity as an 8 on a 1 to 10 scale. This reported outcome is an example of a PRO. The unique information of a PRO, such as the level of experienced pain, the emotional state or frequency of the symptoms occurring in the day-to-day life of out-patients, could be difficult to assess by observing the patient because out-patients tend to visit the hospital at lengthy intervals. This valuable information can help a clinician to adjust treatment plans accordingly [2]. For example, a new medication prescription could be discussed with the patient based on the high level of pain experienced by the patient, which came to the attention of the clinician through the high pain score PRO. PROs are often gathered through questionnaires that are sent out to patients at regular intervals, usually before a planned hospital visit. The resulting PRO is then made available to the relevant clinicians through Electronic Health Record (EHR) programs, where they can navigate the answers to each individual question through a list. These clinicians can use the input of the PRO to learn about the status of the patient's health during clinic preparation, without having to ask all the questionnaire's questions individually during the clinic. Thus, the information provided by a PRO is of great importance and can be used to improve the standard of care provided.

The problem with PROs is that they are often presented in lists of answers to each individual question through large amounts of text, which can be difficult to consume for clinicians. For example, a clinician that must scroll through dozens of questions, vertically listed in textual sentences, to acquire the information that the patient is suffering from pain in their joints may find it too difficult or time consuming to acquire this PRO data. The clinician may then opt to ask the patient directly about the joints where they are experiencing pain. In this case, the data is either not used in the provided care or the clinician is forced to use much of their valuable time to scrolling through the lists of answers. Consequently, clinicians appear repeating questions already gathered in the PRO data, while patients experience their detailed PRO reporting as a time-wasting activity. Many hospitals already use some type of PRO that require their patients to fill in questionnaires regularly. These questionnaires require a significant amount of effort from the patients, making it important that the PRO data resulting from these filled in questionnaires are used well. To ensure that the PRO information is transferred effectively to the

clinician, visualization can provide a solution.

Computer-based *visualization* systems provide visual representations of datasets designed to help people carry out tasks more effectively [5]. Using this definition of visualization, visual representations of PRO data can be designed to help clinicians carry out the providing of care to their patients more effectively. Furthermore, visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods [5]. The interaction between clinicians and patients is vital, and the clinicians are not replaced with computational decision-making. Instead, the human capabilities of the clinicians can be enhanced by designing a *visualization system*. Data visualization has the great benefit of allowing large quantities of data to be interpreted rapidly if it is presented well [6]. Data visualization could therefore be used to improve the difficult to consume format of PRO data that was outlined in the previous paragraph. E.g., these data visualizations in this newly designed visualization system could replace the long lists of textual sentences with PRO results with a concise image that can be consumed quickly by the clinician. The usage of data visualizations may improve the adoption and effectiveness of PROs, but effect of data visualization on PRO adoption is not widely studied. With the increased usage of PROs in clinical research and the increasing interest in data visualization, a significant contribution could be made if PROs could be visualized in way that they are easier to consume compared to the current textual presentation of PRO results.

This study can be the first application of the design study methodology to PRO data. The literature review in section 2 reports existing design studies, including existing design studies of PRO data. This exploration illustrates that there are no design studies that dealt with PRO data yet. This lack of a design study with PRO data means that there has been no methodological characterization of the domain users, or clinicians, that work with PRO data. Section 4 explains the unique benefits of the design study methodology in more detail. With these benefits in mind, a valuable contribution could be made by being the first design study that deals with PRO data successfully. This study could provide important lessons learned for future design studies that deal with building visualizations for clinicians and more specifically, clinicians that work with PRO data.

1.2 Research Questions

The primary goal of this study is to create a validated dashboard with data visualizations of PROs to allow clinicians to provide a better standard of care. Thus, the following research question can be formed to encapsulate this goal:

RQ1: How can data visualizations effectively support clinicians so that they can provide a better standard of care?

Furthermore, this study aims to identify the unique visualization possibilities of the PRO to allow PRO data to be used more effectively by clinicians. Thus, the following research question can be formed to further specify RQ1:

RQ1.1: How can patient reported outcomes be visualized so that they are used more effectively in clinician workflows?

1.3 Thesis Structure

This thesis is structured through 8 main sections, followed by the references and 4 appendixes. Each main section is divided into multiple subsections that each deal with a unique topic. Some of the subsections that present large quantities of information are divided into a third layer of subsections.

The first section has introduced the problem that this thesis will analyze and attempt to solve. A motivation for the importance of this problem was provided, and two research questions have been proposed.

Section 2 initially explains the literature review methodology that was used to create the related work section. Then, the subsequent subsections of section two will go over the four main topics of the literature review: visualizations aimed at supporting clinicians, visualization of multivariate temporal data, patient reported outcome visualization and existing design studies that worked with temporal medical data.

Section 3 provides background on the domain and the domain experts that are involved in this study. Section 3.1 first describes the disease that the patients of the domain experts are dealing with, and the following subsection provides context about the hospital and the team of pediatric rheumatologists that are involved with this study.

Section 4 explains the design study methodology that is adapted in this study and the subsequent subsections go over the methodology behind each of the phases.

Section 5 provides the results of the problem characterization and task abstraction that are part of the ‘discover’ phase. The section presents a model that illustrates the tasks of the studied domain expert, as well as the requirements that have been elicited from interviews with patients and clinicians.

Section 6 explains the transfer of the elicited requirements into a design, from low-fidelity sketches in section 6.1, to a higher-fidelity mock-up design in section 6.2, to a high-fidelity implementation in section 6.3.

Section 7 explains the evaluation of the implementation by first explaining the evaluation methodology in section 7.1, followed by the analysis of the evaluation results in section 7.2.

Section 8 provides a discussion of the strengths and limitations of this study and presents the most important lessons learned for visualization researchers.

The references are presented in order of occurrence in the thesis, using the LNCS format. At the end of the thesis, appendixes are included that provide more detailed versions of the data abstraction, clinician and patient requirements interview protocols and the evaluation protocol.

Abstract	2
1 Introduction	3
1.1 Motivation	3
1.2 Research Questions	5
1.3 Thesis Structure	6
2 Related Work.....	9
2.1 Literature Review Methodology.....	9
2.2 Visualizations to Support Clinicians	11
2.3 Multi-variate Temporal Data Visualization.....	13
2.4 Patient Reported Outcome Visualization	14
2.5 Related Design Studies on Temporal Medical Data.....	15
3 Background	17
3.1 Juvenile Idiopathic Arthritis & Juvenile Arthritis Multidimensional Assessment Report	17
3.2 Domain Description	18
4 Methodology	19
4.1 Precondition phase: Learn, Winnow & Cast	20
4.2 Discover: Problem Characterization & Abstraction.....	21
4.2.1 User Observation	21
4.2.2 Clinician Interview	22
4.2.3 Hierarchical Task Abstraction	25
4.2.4 Patient Interviews	25
4.3 Design: Data Abstraction, Visual Encoding & Interaction	26
4.4 Implement: Prototypes, Tool & Usability	27
4.5 Deploy: Release & Gather Feedback	28
5 Requirement Analysis	29
5.1 From User Observation to Task Abstraction.....	29
5.2 Expert Requirements Interview	31
5.2.1 Clarifying the User Observation: How Do pediatric rheumatologists work?.....	31
5.2.2 How is the JAMAR currently used?.....	32
5.2.3 How can the JAMAR be used more effectively?	32
5.3 Patient Requirements Interviews.....	34
5.4 Elicited Requirements	35
6 Design.....	36
6.1 Visualization Focus Group	36

6.2	Mockup Design	39
6.2.1	Interaction techniques.....	40
6.2.2	Narrative summary	40
6.2.3	Timeline with line charts and medication data.....	41
6.2.4	Rejected visualization techniques.....	42
6.3	Implemented high-fidelity prototype.....	42
6.3.1	Prototype context.....	43
6.3.2	Interaction techniques.....	43
6.3.3	Narrative summary	44
6.3.4	Line charts with textual descriptions.....	44
6.3.5	Deviations from mock-up design	46
7	Evaluation.....	47
7.1	Evaluation Methodology	47
7.1.1	Dashboard tutorial	47
7.1.2	Clinic preparation task.....	48
7.1.3	In-clinic usage task.....	48
7.1.4	Follow-up interview	49
7.2	Thematic Analysis Results	50
7.2.1	Sample Size	50
7.2.2	Codings.....	50
7.2.3	Thematic Mapping.....	53
7.3	Technology adoption.....	54
8	Discussion	56
8.1	Limitations.....	56
8.2	Lessons learned	57
8.3	Conclusion.....	57
8.4	Ethical Considerations.....	58
	References	59
	Appendix A. Data Abstraction.....	62
	Appendix B. Clinician Requirements Interview Protocol	64
	Appendix C. Patient Requirements Interview Protocol.....	67
	Appendix D Clinician Evaluation Protocol.....	68

2 Related Work

The focus of this study is to visualize patient reported outcomes to support clinicians in providing better healthcare to their patients. Therefore, studies that involve the visualization of patient-reported outcomes and medical data in general will be explored. The goal of this section is to gather knowledge of visual encodings, interaction techniques, design guidelines and evaluation methods so that an effective design study can be conducted. Section 2.1 explains the methodology used to conduct the literature review. Section 2.2 will go over the visual encodings, interaction techniques and design guidelines of visualizations created to support clinicians. Then, section 2.3 goes over existing methods of visualizing multivariate temporal data, that are not specific to medical data. Section 2.4 illustrates the current state of patient reported outcome visualization. Lastly, section 2.5 investigates existing design studies that involve the visualization of medical data to learn about the evaluation methods used.

2.1 Literature Review Methodology

In this section, the sources used for the literature study are outlined and for each source, a description of the elimination process of papers is given. The sources used are the TimeViz Browser [7], Google Scholar, the snowballing technique and using existing systematic reviews. Figure 2.1 illustrates the number of papers retrieved from the TimeViz, Google Scholar and snowballing sources. The rest of this section will provide more elaboration on the inclusion process and the queries used to retrieve papers.

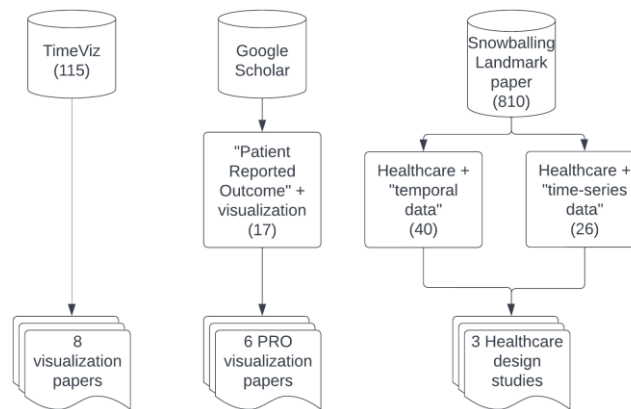


Figure 2.1 Paper exclusion process model

Source1: The TimeViz Browser based on Visualization of Time-Oriented Data [7]

PRO data is generally acquired through questionnaires on a regular basis. The result of this data collection is a variety of data points of multiple variables over time. All data related to time can be referred to as temporal data [5]. The data collection of PRO data occurs sequentially, with new data points coming in as time progresses. Thus, PRO data can be referred to as time-series data which is an ordered sequence of time-value pairs [5]. The TimeViz web-based browser displays 115 visualization techniques for temporal data and was last updated in 2015. All the visualization techniques that are displayed in the TimeViz browser originate from a published scientific paper, making it an excellent resource to find scientifically validated visualization techniques.

The browser uses the term ‘time-oriented’ for same given definition of temporal data. Relevant visualization techniques have been selected based on their ability to visualize multiple variables in a concise image, visualizations related to medical data have been given priority. Visualization types that did not have any studies related to medical data, such as the circular visualizations, have also been included in the analysis. In total, 8 visualization techniques have been included in the literature review of temporal data visualization.

Source 2: Google Scholar

The following search queries have been used to identify papers: information visualization healthcare, patient reported outcomes visualization. Through the patient reported outcomes visualization search query, 17 papers could be identified that dealt with data visualization of PRO data in the first 5 pages of search results. Of the identified papers, 6 were included in the literature review for their relevance to presenting PRO data more effectively through data visualization. The excluded papers did not provide any useful insights on the effective visualization of PRO data.

Source 3: Snowballing

Citations of systematic reviews were used to identify relevant papers. Furthermore, the ‘cited by’ of key papers in visualization literature were used to identify more recent state-of-the-art literature. Design studies have been identified by using the snowballing technique to find studies that have cited the ‘9-step framework’ presented in the ‘design study methodology’ paper by Sedlmair et al. [8]. The landmark study has been cited 810 times. These 810 papers have been filtered on a relation to healthcare and temporal data which resulted in 40 papers. Moreover, 26 papers were identified that dealt with time-series healthcare data specifically. Out of the identified papers, only the papers that used the 9-step framework were included and considered design studies. Studies that follow the 9-step framework use the same methodology that is used in this thesis, which allows the execution of each step to be learned from so that this thesis can conduct a better design study. Some papers only referred to the 9-step framework as a potential improvement for their study [9], but did properly use the 9-step framework. No design studies could be identified that dealt with patient reported outcomes.

Source 4: Systematic reviews

Landmark systematic reviews are used to identify visualization papers by using the citations of these papers with snowballing. For example, West, Borland & Hammond [10] identified visualization techniques used with Electronic Health Record (EHR) data. Their review is limited to articles published between 1996 and 2013 and included only studies with visualizations other than standard graphs such as bar charts, pie charts or line graphs. A review of non-standard visualizations allows us to identify novel visualization techniques that are relevant to the medical domain. Thus, this systematic review was used to identify landmark visualization papers in the medical domain.

2.2 Visualizations to Support Clinicians

This section analyzes existing visualizations aimed at supporting clinicians to learn about effective visualization and evaluation methods. Meta-analyses have identified effective visualization and interaction techniques to support the analysis of Electronic Health Record (EHR) data by medical professionals. EHR data contains data about patients and this data is used by clinicians to gain a better understanding of their patients to provide better treatment. Visualizations of EHR data is studied because it involves the same target users as PRO data, namely clinicians. West, Borland & Hammond [10] analyzed 18 studies involving visualization of EHR data and found that with the large amount of data of EHR, it can be difficult to identify meaningful patterns and that this can be mitigated by using tools such as zoom, pan and filter. This is in correspondence with the problem of PRO data being too time consuming for clinicians. The color, density and filtering techniques are commonly used to distinguish variables or temporal events [10]. E.g., a user could filter out a specific disease symptom such as ‘pain’ and distinguish the experienced level of pain with red color indications. Presenting a great deal of information on a single screen where the user can interactively explore the information is an important design feature to medical professionals. Thus, a tool developed for the clinician user group can present large quantities of information if the right interaction techniques are supported.

To get an understanding of the strengths and limitations of existing visualizations used by the clinicians practicing at a large number of hospitals, the most cited visualization system LifeLines [11] and its successor are analyzed. One of the earliest (1998) attempts at visualizing EHR data represents the application of the color, density and filtering techniques well while showing how this visualization can be achieved in a concise image. LifeLines [11] summarizes EHR data as a set of lines and events on a zoom-able timeline. The display visualizes problems, allergies, diagnoses, complaints, and medications of the past 6 months in one screen. The Java user interface allowed its users to click on any of the words to gain more detailed information. The most critical and prevalent aspects such as the main diagnosis were colored in red. The

visualization illustrates a method to display many variables on a single timeline, simultaneously, while not overloading the user with information. This is achieved by stacking horizontal timelines that indicate the duration of various diseases and disease symptoms. Furthermore, changes in medications are indicated in the bottom half of the image, allowing the clinician to cross-reference the changes in medication with the disease symptoms. LifeLines was extended with LifeLines2 that visualized the temporal EHR data of multiple patients, instead of a single patient. LifeLines2 used colored multiple-colored triangles to visualize the duration of diagnoses instead of the continuous horizontal lines. Follow-up research showed that the horizontal lines can give misleading indications of the duration of symptoms because it displays a continuous line while disease symptoms vary over time. LifeLines2 has been validated through qualitative interviews and quantitative measurements of task completion times and task errors. The evaluation showed that the interface can be learned quickly and that it addresses the need to rapidly review results and spot patterns of interest. The authors mention that the line used to denote the duration of diagnoses in the original LifeLines wrongfully gives the idea of the real disease duration, [12] while the triangles are insufficient [13]. Therefore, using visualizations with datapoints that are unconnected by lines may provide a more accurate representation of the progression of a disease. While both LifeLines and LifeLines2 manage to visualize multiple disease symptoms over time, they are both limited in their ability to distinguish between the severity of symptoms. Visualizing disease symptoms as many separated events on a timeline improves the accuracy over a continuous line but fails to visualize the varying symptom severity over time.

Since there is a clear need for a better representation of the duration and severity of the disease activity on a timeline, several more recent EHR visualizations are analyzed. The second most cited EHR visualization is VISITORS [14] that is based on KNAVE [15] and KNAVE-II [16]. VISITORS presents similar interaction techniques to LifeLines: a large overview of the data, zooming and panning, task filtering and details on demand. These techniques allow the users to select multiple relevant variables to visualize these variables with multiple line graphs. The time on the x-axis can be adapted by panning left to right and the scale can change by zooming, appropriately matching the need for a temporal view. However, the large number of line graphs in a single image reduces the clarity of the overview, calling for a simpler solution.

Data dashboards are commonly used in the medical domain and have been shown to improve the care provided by medical professionals [12]. A review of studies of dashboards for improving patient care by Dowding et al. in 2015 was able to identify 11 studies that evaluated the impact of the dashboard on desired outcomes or the perception of clinicians of the utility of such a dashboard. 9 of the 11 studies identified by Dowding et al. used color coding in the format of a traffic light approach where green indicates that there is no action to be taken by the individual and red indicates that an action is required [17]. Thus, medical professionals are likely to be comfortable with the use of these traffic light colors for different levels of required attention. E.g., the reporting of a disease symptom in red could gain the attention of a clinician while it may otherwise have been overlooked. Another visualization system that used color is CareCruiser [18] which visualized steps of a treatment plan with a flow chart and the hierarchical structure of treatment plans tree graph. The effects of a treatment plan on oxygen saturation levels are visualized for multiple patients simultaneously by stacking multiple dot plots vertically. Deviations from the intended value are visualized with gradations of magenta

within each dot plot. This approach allows the user to see information of multiple patients in one overview, while calling attention to strong deviations from the intended values with colors. Both the usage of traffic light colors and the magenta gradients appear to be effective in visualizations for clinicians, which will be further investigated in this thesis. Besides the use of color, related studies have given an indication of the most important requirements for visualizations aimed to support clinicians. Kosara & Miksch [19] surveyed visualization techniques in the medical domain and that found intuitiveness, focus & context of time/data, combination of values, seeing developments, finding patterns and discovering intervals were the most important requirements. Any proposed visualization technique should attempt to adhere to these requirements to maximize effectiveness, while the use of traffic light or magenta gradient colors could further improve the effectiveness of visualizations for clinicians.

This section has analyzed a variety of visualization systems that are made to support clinicians. Existing research suggests that clinicians are capable of consuming large amounts of data in a single screen, if the appropriate interaction techniques such as filtering, and zooming are used. Traffic light colors and magenta gradients are proven methods to navigate the attention of clinicians, both color types could be further investigated. There is a need for an effective visualization of multiple disease symptoms over time, where the severity of the symptoms can be distinguished. The next section will analyze multi-variate temporal data visualization to find a suitable technique for this.

2.3 Multi-variate Temporal Data Visualization

This section will go over various methods of visualizing multiple variables over time. The section analyzes multiple visualization techniques involving stacking horizontal timelines and circular views to find the most appropriate and accurate method of visualizing multi-variate temporal data. This is important because the PRO data that is studied, contains multiple variables that need to be visualized over time. Aigner et al. executed a systematic view of time-oriented data where they discuss the danger of misrepresenting what happens between data points in a visualization [20]. Thus, it is important to identify visualization techniques that do not misrepresent data about points of time where there is no information. TimeSearcher 2 [21] enables users to select parts of a timeline with click and drag interactions to receive exact values for the selected timeframe. Multiple timelines can be stacked vertically, and the timelines are always presented with detailed information in a secondary window. IDMVIS [22] visualized temporal event sequences by presenting a timeline for each day and stacking multiple timelines vertically. Individual days could be selected to view a more accentuated and vertically stretched view of that day's temporal event sequence. IDMVIS uses multiple horizontal timelines that each represent a day. Both TimeSearcher 2 and IDMVIS show how multiple timelines can be stacked in a concise presentation. The approach taken in TimeSearcher 2 allows for the acquisition of more detailed information compared to IDMVIS, which could avoid the danger of misrepresentation of the data.

Many visualization techniques for temporal data involve a circular view to visualize large amounts of data in a compact format. CircleView [23] divides a circle into 6 slices that each

present an attribute, and each attribute is sliced into 10 parts that represent different timeslots with the timeline starting at the center of the circle. A green color scheme is used to visualize the attribute values and red is used for values that surpassed a threshold. This approach allows for visualization of large quantities of attributes over time without overloading on information but is limited to only 10 timeslots and the exact time of data points is unclear. Enhanced Interactive Spiral [24] turns a timeline into a spiral with two-tone color coding that indicate values with customization options to adjust the color schemes and to move through time. This approach effectively visualizes large quantities of data points for a single attribute in a more compact figure than a horizontal timeline. However, while the visualization is compact, the exact dates of each data point only become apparent after using the hovering interaction technique, unlike the identified horizontal timelines.

This section showed that the visualization of many variables over time can be achieved through both the stacking of horizontal timelines and the use of circular displays. However, the use of horizontal timelines is the more validated approach that is also validated to work well with the clinicians as target user.

2.4 Patient Reported Outcome Visualization

This section illustrates the current state of visualization research that involves PRO data to learn how PRO data can be visualized. Research that visualizes the PRO to the patient is discussed, followed by existing research of PRO visualization for clinicians. For the PRO in pediatric rheumatology specifically, there is a push to acquire and deliver results to patients electronically [25], such as PROMIS [26] that incorporated patients in the collection of PRO through their own smartphones. Visualizing the medical data of patients directly to the patients require additional attention due to the highly personal nature of the data presented which can influence the emotional state of the patient. Negative emotions lead to a decrease in performance in graphical perception tasks requiring visual judgment [27]. Visualizations with prostate cancer patients reported not being able to process any information after hearing their severe diagnosis due to heightened negative emotional states [28]. Patients have unique preferences when it comes to the visualization of their own PRO data. Stonbaker, Porass & Schnall conducted an experiment to find the preferred visualization of longitudinal PRO data [29]. During the thinking-aloud survey about a patient report, patients mentioned a preference for bar charts over line graphs and scatterplots. Patients experienced traditional line charts and bar charts as impersonal, and it made them feel like they are looking at stock market graphs. The inclusion of emoticons, ranging from sad to happy, in the traditional visualizations were favored by the patients. The inclusion of emoticons also increased comprehension of the information presented, the bar graph with emoticons added to each bar was identified as the most preferred and most comprehensible visualization [29]. This study illustrates the need to include patients in the design process of the visualization design. Engaging patients in the PRO collection also improves symptom reporting and illness understanding. PROACT [28] visualized PRO data to patients over the age of 65 by using a narrative style and simple visualizations such as bar charts and pie charts. Their study found that their older patients had low visual literacy and could were not able to use interaction techniques such as sliders effectively. Removing all interaction and

presenting more information in the form of pie charts in a single view was found to be more effective. Both patients and clinicians prefer the presentation of more information in a single view. However, patients have unique preferences when it comes to the inclusion of emotion in charts and are unable to use interaction techniques effectively. Therefore, it is important to include the patients in the design process of the visualization design so that both the needs of the clinicians and the patients can be met.

Specific filtering options while keeping the visualizations in view are a validated approach to visualize PRO data. PROBE [30] is a dashboard created with Power BI that visualizes the resulting percentages of the patient population of PRO with pie charts and bar charts. The dataset can be filtered in the dashboard with several personal attributes, health related attributes and disease specific attributes while always keeping the created visualizations in view. While the visualizations used are simple in nature, the dashboard succeeds in providing a wide array of filtering options to specify results. To conclude, the visualization of PRO has unique requirements because of the personal nature of the patient data and further research is necessary to identify these. This section has analyzed studies that deal with the visualization of PRO data. However, no studies could be identified that use the design study methodology on PRO data. That means that there is no clear characterization of the problems faced by the clinicians and patients that are involved with PRO data. While no design studies could be identified that studied the visualization of PRO data, it is still important to analyze design studies that deal with temporal medical data to learn about the visualization and evaluation techniques used in design studies that dealt with similar data.

2.5 Related Design Studies on Temporal Medical Data

Design studies that deal with temporal medical data are analyzed to use the lessons learned to allow for a better conducting of a design study with this data. Primarily, the evaluations of the designs are analyzed to learn about the evaluation methods used. Moreover, the key lessons learned during the design studies are considered. Existing design studies have been successful at visualizing medical data with multiple superimposed windows. IDMVIs [22] proved that superimposed distinct windows may be useful for treating chronic diseases that require intensive self-monitoring and treatment adjustment. The solution was evaluated by letting clinicians explore data through traditional methods and then do the same with the IDMVIS solution afterwards. Thinking aloud was requested during the evaluation task to acquire the opinions of end-users throughout the experiment. Their evaluation illustrates how IDMVIS correctly reflects clinician workflows and allows the clinicians to accurately identify patterns across timelines. IDMVIs [22] used a rigorous evaluation method to achieve this, the inductive qualitative analysis. They did the coding across three authors, using more authors for coding themes leads to more accurate evaluation results.

QualDash [31] is an adaptable generation of visualization dashboards for healthcare quality improvement that also uses multiple windows that are sequentially introduced. These studies

indicate that clinicians can effectively use a dashboard with many visualizations in a single interface. Qualdash applied their research at 5 different hospitals [31], which increases the representability of the results. Qualdash evaluated the perceived usefulness through a case study. The analysis was not explained in a rigorous manner through an established method such as the thematic analysis. The design study conducted to create ConVIScope suggests that their clinician participants wanted to be enabled more to compare various metrics [32]. However, their current design did not allow for comparison sufficiently. Considering the earlier findings that clinicians seem to be exceptionally able to consume large quantities of visualizations in a single interface, it seems that the comparison of various attributes can be enabled by visualizing multiple metrics in a single interface. The evaluation done in ConVIScope [32] does also not use a rigorous methodology, but instead lists general impressions. This analysis of existing design studies on temporal medical data shows a theme in a recommendation of presenting large quantities of information in a single interface for this user group. Qualitative evaluation methods like case studies at multiple hospitals or the inductive qualitative analysis with multiple coding authors have been used to execute rigorous evaluations.

The related work section has analyzed existing visualization systems created for clinicians that outlined the unique characteristics of clinicians and identified the need for an effective visualization technique of multiple disease symptoms over time, where the severity of the symptoms can be distinguished. The primary identified techniques to achieve this are the vertical stacking of horizontal timelines and the use of circular views, of which timelines are superior in accuracy. Patient reported outcome visualization has been studied, which showed the unique characteristics of patients and the importance of including patients in the design of a visualization system with PRO data. Lastly, existing design studies have been analyzed to find evaluation techniques and key lessons learned.

3 Background

The previous section concluded that it is necessary to study both the clinicians and patients with the visualization of PRO data. To investigate how data visualization can effectively support clinicians to provide a better standard of care, a case study is conducted with pediatric rheumatologists and their patients. Section 3.1 provides general background information on the diseases that these clinicians treat as well as the associated PRO data that is reported by their patients. Section 3.2 continues with an explanation of characteristics that are specific to the studied hospital and the pediatric rheumatology department.

3.1 Juvenile Idiopathic Arthritis & Juvenile Arthritis Multidimensional Assessment Report

Juvenile idiopathic arthritis (JIA) is a term for diseases that encompasses all forms of arthritis that begin before a patient is aged 16 years that persist for more than 6 weeks and are of unknown origin [33]. It is the most common chronic childhood auto-immune disorder that affects 7.8 patients per 100.000 [34]. There are disease categories of JIA that each have distinct methods of presentation, clinical signs and symptoms, mainly based on disease severity and the involvement of other organs than the joints. The cause of the disease is poorly understood but seems to be related to both genetic and environmental factors, which results in the heterogeneity of the illness [35]. If left untreated, JIA can be a disease resulting in life-long disability [33].

Traditional assessment of children with JIA has been on strict outcomes such as disease activity (e.g., number of inflamed joints), disease remission, joint damage, and organ system damage. These outcomes are all quantified by the treating pediatric rheumatologist. However, other aspects can also affect the child's disease outcome and quality of life such as physical, social, emotional, intellectual and economic aspects [35]. These aspects can be measured through parent/patient-reported outcomes (PRO) that reflect on the parents' and children's perception of the disease course and effectiveness of therapeutic interventions.

The Juvenile Arthritis Multidimensional Assessment Report (JAMAR) is a questionnaire designed to acquire these PROs through regular administration in daily clinical practice. The JAMAR measures overall well-being, pain, functional status, morning stiffness and overall level of disease activity, rating of disease status and course, proxy- or self-assessment of joint involvement and extra-articular symptoms, description of side effects of medications, and assessment of therapeutic compliance and satisfaction with the outcome [36]. These measures are all patient reported.

The JAMAR is widely used in clinical practice in many hospitals and is available in many languages. For example, the JAMAR has been translated to the Dutch language and has been validated as a useful tool to assess children with JIA and was deemed suitable for both routine clinical practice and clinical research [37]. In most languages the JAMAR has a version for younger children which is filled out by the parents, and a version for older patients who can do this themselves. In the UMCU, the JAMAR is filled out by each patient or parent prior to the

outpatient clinic visit since 2011. The JAMAR results in a single dataset that is ready for quantitative analysis.

3.2 Domain Description

The project is in collaboration with clinicians and researchers of the department of pediatric immunology and rheumatology at a Dutch hospital. The project involves medical data of the patients collected from electronic medical records. The department consists of five pediatric rheumatologists who see a total of 400-500 unique patients each year. The project was initiated by one pediatric rheumatologist and two researchers at the hospital. The primary condition of the patients is JIA.

The department stores a variety of data on its research data platform (RDP) that is available for analysis by data analysts. The RDP is a weekly pull of data from amongst others the electronic medical record and possess information about patient demographics, clinic dates, diagnoses, medications and lab results. This dataset is anonymized with pseudo-ID numbers and is therefore ready for analysis. Data of roughly 1000 to 1500 patients is stored of children with JIA aged up to 18. For this project, the most relevant datasets are the patient JAMAR and parent JAMAR results. An abstraction of the JAMAR Patient dataset can be found in table 1 of Appendix A.

The results of the 15 questions of the JAMAR are currently presented as text and require too much time investment for clinicians to analyze during each clinic. The currently used EHR system makes it difficult to view the progression of the results of the JAMAR over time for a single patient. Furthermore, it is not possible to gain a general overview of the progress of the JAMAR results over time for patient groups or the patient population. The JAMAR contains unique information that could support decisions of treatments for JIA, so it is essential that this information is presented properly so that it can be used effectively during clinics to improve the standard of care. Besides improving the standard of care, an effective presentation of the results of the JAMAR would also provide patients and their parents with feedback on the filled in questionnaire over time. This feedback could motivate patients and parents to keep investing time and effort into filling in the questionnaire, resulting in a better adaptation of the JAMAR.

The department currently uses a EHR interface that displays medical data in a tabular format. Clinicians use this EHR system before, during and after their 20-minute clinics to manage patient data. The system displays the results of the JAMAR by listing the scores that have been derived from multiple question results such as the ‘Score_ADL’ in Table 1 of Appendix A, as well as textual values to answers of each question. All these values are displayed in a single text box and summarize the data for a single patient and for one data entry. The display of the results in a single textbox that requires the clinician to scroll through a long list of sentences. This perfectly illustrates the problem posed in the introduction of this thesis. To tackle this poor presentation of PRO results at this hospital, the design study methodology is used which is detailed in the next section.

4 Methodology

This section will outline the methodology used to conduct this study. The methodology is an adaptation of the well-established design study methodology. The design study methodology is a form of problem-driven research, where the goal is to work with real users to solve their real-world problems. A *design study* is defined as “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.” [8] The nested model for visualization creation proposes four nested layers namely domain problem characterization, data/operation abstraction design, encoding/interaction technique design & algorithm design [38].

To conduct the design study, the 9-step framework by Sedlmair, Meyer & Munzner [8] has been adapted to this project domain. This 9-step framework is illustrated in Figure 4.1 which lists the 9 steps in sequential order while also including a connection to all previous steps, from each step, which indicates that previous steps can be revisited at any time during the study. The figure categorizes the steps into three phases which will be explained throughout this section. Section 4.1 explains the steps taken during the precondition phase. Section 4.2 describes the discover step. Section 4.3 will explain how the design is created. Section 4.4 describes the implementation based on the created design. Lastly, section 4.5 explains how the implementation is deployed. The reflection can be found in the discussion in section 8.

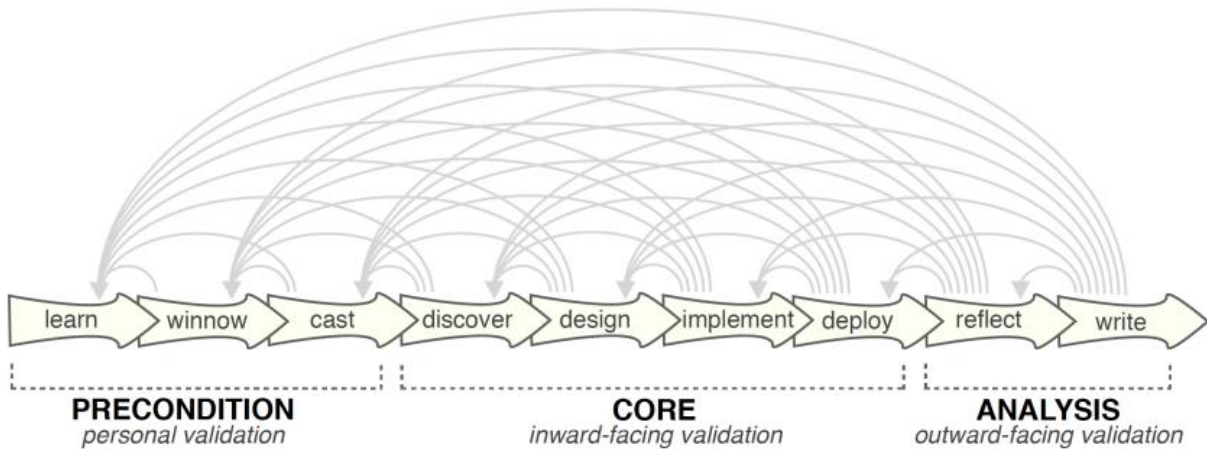


Figure 4.1 9-step framework [8]

The application of this 9-step framework can lead to three types of research contributions. The first contribution of this project will be a characterization of the domain problem of improving

decision making of pediatric rheumatologists to improve understanding between visualization experts and healthcare professionals. The second possible contribution mentioned by Sedlmair, Meyer & Munzer is a validated visualization design. This project created a dashboard with information visualizations that is validated through a study involving pediatric rheumatologists as end-users. Thus, it is likely that a significant scientific contribution was made here. The third type of contribution mentioned by Sedlmair, Meyer & Munzer is the reflection on the design study and its retrospective analysis in comparison to other related work. The result of this reflection can lead to improved guidelines for conducting design studies in future research. This reflection can be found in the discussion in section 8.

4.1 Precondition phase: Learn, Winnow & Cast

This section explains the execution of the first phase of the 9-step framework, the pre-condition phase. The first step of the pre-condition phase is to (1) Learn which involves the familiarization with visualization literature, including visual encoding and interaction techniques, design guidelines, and evaluation methods to conduct an effective design study [8]. This step has been executed in section 2 in the form of a literature review.

The second step of the 9-step framework, (2) Winnow, involves finding the correct collaborations for the projects. It also involves asking essential questions before the project regarding the availability of data and ethical considerations. Weekly meetings were conducted early in the project to find possible collaborations and to ensure that the data is readily available for the researcher in an anonymized format. The selection and identification of collaborators was achieved through frequent meetings in the early stage of the project with daily supervisors at the hospital that could take on the role of connectors. Through the connectors, connections could be made to fellow tool builders, that will maintain the deployed dashboard after this project, and the end-users of the dashboard. The connectors also provided access to the data, the current EHR software and a fully anonymized version of the dataset that can be shared with third parties for evaluation. Access to the data was provided early in the project, allowing for the creation of a data abstraction during the exploration of visualization literature.

The third step of the 9-step framework is (3) Cast, where the most promising collaborations that were found through early meetings are cast into roles. The following roles were cast for this project:

- *Researcher* that conducts the research as part of a master thesis, supervised by two supervisor researchers from Utrecht University. The researcher is responsible for setting up interviews, experiments, creating the dashboards with data visualizations and writing the majority of the research thesis and paper.
- *Daily supervisors* at the hospital that is the domain of study for this project. The most important role of these supervisors is the role of connector, connecting the researchers to domain experts, peer students and fellow tool builders.

- *Clinicians* are the domain experts and end-users of this project. They provide the requirements in interviews and are used for the evaluation of the designed artifact. The clinicians also provide access to patients.
- *Hospital Team Data Solutions* are the fellow tool builders that will maintain the designed artifact at the end of the project.
- *Peer students* can be asked questions during the project, and they can be involved for the usability study of the developed dashboard. These peer students generally have medical backgrounds and are authorized to view patient data.

4.2 Discover: Problem Characterization & Abstraction

After the pre-conditioning phase is completed, the core phase of the design study is initiated with the (4) Discover step, that aims to characterize the problem and aims to create a task abstraction. To better understand the problem, section 4.2.1 explains methodology of the observation of the target user during their clinics with patients. Section 4.2.2 explains how this user observation was followed up with an interview with the clinician. Section 4.2.3 describes how a hierarchical task abstraction is used to illustrate the findings identified tasks of the target user. Lastly, section 4.2.4 describes how patients are interviewed to understand their point of view. The results of the discover step can be found in section 5.

4.2.1 User Observation

To characterize the problem and come to a task abstraction, a user observation with the *fly-on-the-wall* [39] method was performed. Using the fly-on-the-wall method means that the researcher is present during the daily activities of the target user, but does not interfere in a significant way, aiming to gain an objective view of the daily activities of the target user. Two 4-hour clinics that each involved around 8 patients were silently observed. A ‘clinic’ refers to a 20-to-40-minute meeting between a patient and their doctor where the disease and its treatment are discussed. The first clinic was used to get a general impression of the domain early in the project. The second observed clinic was used to get a more concrete description of all tasks performed by the clinician. All observed tasks were noted down with pen and paper to gain an insight into the tasks performed. Because of the confidential nature of an out-patient clinic, complete absence of the researcher that is required for fly-on-the-wall observation was not possible. Complete absence of the researcher would require the clinics with patients to be recorded, which is a breach of confidentiality. The researcher made introductions to the patients to inform about the purpose and to ask for permission to be present and because of the limited space of a clinic room, the researcher is always in view. However, to limit the implications the presence of the researcher has on the results of the characterizations, the researcher was seated as far away from the patients as possible, next to the clinician while remaining out of the sight of the clinician for most of the clinic. This method resulted in an as accurate as possible

indication of the tasks performed by the clinician given the circumstances, and the observation is followed up with an interview to further solidify these results.

4.2.2 Clinician Interview

To specify the domain problem and validate the observations made during the user observation, a *semi-structured interview* [40] was conducted with the observed clinician 2 days after the second observation. The *critical incident* [41] technique was used by referring to specific observations and by letting the clinician narrate their rationale behind their actions and decisions. This section lists the 11 core questions that are asked during the interview and provides a description and rationale for each question. The complete list of interview questions, with the detailed sub-questions, can be found in Appendix A. For the first question Q1 a specific observation is pointed out, to then ask if that is how the JAMAR results (described in section 3.1) are typically used. Follow-up questions were asked in cases where the response lacked detail. These follow-up questions can be found in Appendix A.

Q1: How do you typically use the results of the JAMAR in your clinical practice?

Furthermore, the clinician was asked about previous attempts at tackling the poor presentation of PRO results to learn about potential shortcomings of these previous attempts to visualize the PRO data. This question also aims to learn what previous attempts did well according to the clinician, so that the shortcomings and strong characteristics can be used to create a better suited system.

Q2: Have there been systems used in the past to tackle potential limits of the current HiX system?

During the second half of the interview, the clinician was presented with the JAMAR questionnaire and asked to narrate where they see value in the results of the questionnaire. The clinician was tasked to create a priority list out of the 15 JAMAR questions to identify the most valuable questions, allowing for the creation of visualizations that align with the priorities of the target users. The aim of this question is to get an understanding of the most valuable data to prioritize their representation in the visualization design.

Q3: Out of the 15 main questions of the JAMAR which answer would you consider the

most valuable in treating patients with JIA?

Q3.1: Could you arrange the five most important JAMAR questions?

Q3.2: Is there any aspect of the JAMAR questionnaire you do not value or intend to utilize while providing care?

Then, a question is asked to find essential functionality of the to be developed solution and to learn about how the visualizations can be used during the clinic. The clinician is asked to narrate what they would like to be able to do with the PRO data, that they are currently not empowered to do. They are told to not limit themselves by any practical or technological limitations. The result of this question gives an insight into the ideal use case of the PRO data from the point of view of the clinician.

Q4: What would you like to be able to with these JAMAR results that you currently do not feel empowered to do?

Next, the use case of the newly developed system is further clarified by asked the clinician how they would like to use the system in Q5. The optional follow-up questions inquire about using the system to learn data themselves versus using the system to present data to the patient. This question aims to acquire insights on how the clinician envisions their preferred use case of the system.

Q5: If a new system were to be build and implemented to better visualize the JAMAR, how would you like to be able to use this system?

After the clinician was given the opportunity to freely narrate their ideas, desired functionalities and the use case of the newly developed visualization system, more specific questions are asked to create a more concrete requirements specification. Q6 to Q9 are aimed to get an understanding of the priority of the most important data visualization functionalities. These questions are also expanded with any additional functionalities that have come up out of Q4 and Q6. For each of these questions, follow-up questions are asked if necessary to get an as concrete as possible answer.

Q6: How important is it for you to be able to enter new information to this new system?

Q7: To what extend would you like to be able to search for patients with this new system?

Q8: How important would it be for you to identify patients with outlying values?

Q9: How important would it be for you to be able to compare patients on measurements?

Since a potential use-case involves the display of visualization to child patients, the clinician is directly asked about potential ethical dilemmas that this use case may bring. Q10 aims to clarify if all PRO data can be safely displayed to the patient, or whether there is data or combinations of data that are not fit for display to the patient.

Q10: Is there any information derived from the results of the JAMAR that is not appropriate to be shown to the patient?

Lastly, the final question is asked that aims to clarify any ideas of the clinicians that did not yet come to attention during the interview. To give the clinician some inspiration, predictive analytics and other data sets were suggested to aid in the answer given. This question should minimize the chance of any valuable ideas regarding the developed system lingering in the head of the clinician.

Q11: Are there any other elements that we have not discussed that you would like to see visualized in the new system?

Q11.1: Would you value predictions of future developments in the new system?

Q11.2: Are there elements besides the JAMAR results that you would like to see visualized?

The answers to these questions have been transcribed and transformed into the task abstraction and requirements. The report of the answers can be found in section 5.2.

4.2.3 Hierarchical Task Abstraction

To abstract the observed actions by the clinician, a *hierarchical task abstraction* [42] was created. The purpose of the task abstraction is to learn how visualizations can intervene in the existing tasks. This task abstraction is created by listing the observed tasks in sequential order, with more complex tasks having more concrete subtasks that have occurred in some of the observations. Using hierarchy allows for the creation of a model that is accurate, yet concise. For each of the tasks, the number of occurrences is denoted in the model. This should result in a model that accurately represents the workflow of the clinician so that the design can be fitted to the right real-world tasks.

4.2.4 Patient Interviews

During the interview with the clinician, it became apparent that the patients are also an essential part of the target users of the visualizations. The patients at this hospital have a full overview of what is done on the 2 monitors that the clinician handles during the clinic. This means that the patient can see how their JAMAR results are represented on the screen of the clinician during a clinic, making them an essential part of the target users. Thus, patient interviews were performed by accompanying an out-patient clinic for the third time and by asking patients for a short interview to improve the standard of care of the hospital. Patients and their parents were both asked to provide informed consent. The clinician provided consent for the interviews beforehand. The main benefit of interviewing patients right after their clinics is the fact that most patients will have filled in the JAMAR just before the clinic, meaning their experience of filling in the JAMAR is fresh on their mind. During the interview the critical incident technique is used again by asking patients to narrate their experience with filling in the JAMAR questions in Q1. The patient is presented with a printed version of the JAMAR questionnaire and asked to narrate their experience with the JAMAR questionnaire. By presenting the JAMAR questionnaire, the patient should be reminded of the contents of the questionnaire to enable them to give more detailed answers.

Q1: How would you describe your experience with filling in the JAMAR questionnaire?

Q1.1: For how long have you been filling in the JAMAR?

After learning about the experience of the patients with the questionnaire, the patient is inquired about potential confusing aspects of the questionnaire. This question is aimed to ensure that the patient fully understands the questionnaire and that any confusing aspects are clarified. This will allow the patient to answer the third and most important question optimally.

Q2: What elements of the JAMAR questionnaire were confusing to you?

Then, the patient is introduced to the idea of a doctor that has all patient data available in his memory and has analyzed all the patient data, he is aware of all your medical data throughout the past. Besides the patient data, this doctor knows of all the patient data of the patient population at the clinic and the doctor has a computer in front of him and all the possibilities of technology available to him. This hypothetical super doctor is used as a simplification of a data visualization system. After this idea is well established in the patient, they are asked what they would like to discuss or learn from this doctor in Q3. Follow-up questions are asked based to the patients that require more direction or assistance.

Q3: Given these limitless possibilities, what would be interesting for you to discuss or learn from your doctor?

Q3.1: Would you value seeing your answers over time? If so, why?

Q3.2: See values that are very different from normal? If so, why?

Q3.3: Would you value seeing how you answer in comparison to the patient population? If so, why?

The first two questions that were presented mostly serve to prime the patient to answer the third and most important question that identifies what the patient would like to learn from a visualization system. Based on the elicited requirements from the user observation, clinician interview and patient interviews, a design is created. The methodology behind the design step is outlined in the next section.

4.3 Design: Data Abstraction, Visual Encoding & Interaction

The design (5) step is where the requirements and task abstraction are turned into visual encodings and interaction techniques. This is achieved by first creating an abstraction of the dataset that is available, to achieve a good understanding of the data. Upon reaching data understanding, visual encodings and interaction techniques can be created.

Access to the dataset was provided early in the project, allowing for an early data abstraction which will be described here. By doing the data abstraction early, most of the literature review about state-of-the-art visualization techniques could be conducted with the data types in mind. The data abstraction is created in a table by listing the attribute name, data type and example values for each of the attributes in the JAMAR patient dataset; this table is presented in

Appendix A. Only the JAMAR patient dataset is abstracted since the JAMAR parent dataset contains the same results, but the results were gathered through slightly different questions to accommodate the third person perspective of a parent. The data abstraction shows how the dataset consist of roughly equal number of categorical, ordinal, and Boolean values. There are 8 categorical attributes, 8 ordinal attributes, 8 Boolean attributes and 5 attributes with natural text. Two categorical attributes in the data abstraction table consist of sub-variables. V1-V15 consists of 15 categorical sub attributes, while V119-V128 consists of 10 categorical sub attributes. Given this data abstraction it can be concluded that the dataset consists of primarily categorical, ordinal, and Boolean values. However, when sub variables are considered, there is a clear need to find visual encodings that can visualize a large quantity of categorical attributes over time.

Based on the data abstraction and the results of the discovery phase, the visual encodings and interaction techniques are identified. To identify good visual encodings and interaction designs, Sedlmair, Meyer & Munzner suggest a broad consideration space, which is especially vital to problem-driven research [8] such as this study. The visualization techniques that have been identified through the literature review were filtered down to a narrow proposal space through paper prototypes. Furthermore, sketches were made of new types of visualizations that are independent of technology based on the JAMAR data. These propositions were presented to clinicians and peer students during a visualization focus group to identify several good solutions to visualize the JAMAR results. Section 6.1 contains the results of the visualization focus group, as well as the visualization sketches. Paper prototypes and mockups were altered based on the feedback gathered during research meetings so that a suitable design could be identified through this iterative process. The higher-fidelity mock-ups are presented in section 6.2. Once the design was complete, the development of the high-fidelity prototype started. The next section details how these prototypes are created and how the tool is implemented.

4.4 Implement: Prototypes, Tool & Usability

During the implement (6) step of the 9-step framework, a high-fidelity prototype was developed based on the created design. A common pitfall of design studies is investing too much time in coding a solution, that the developer does not want to throw it away when it is not the right solution [8]. Thus, rapid-prototyping was essential by following the lessons from agile development [43]. The hospital data analysis infrastructure of the domain of study is based on the R programming language. The data analytics team that was identified as fellow tool builder that will maintain the deployed visualizations after the project was finalized, was also most familiar with the R language. Thus, the creation of visualizations of the R language was preferred for development. However, visualization algorithms based on the D3.js and Python programming languages were also explored to not exclude a potential other solution. The visualizations created through these 3 languages were coded as quickly as possible so that they can be evaluated and either altered or thrown away based on the feedback from target users during the research meetings. To accommodate for the confidentiality of the patient data, a version of the tool that is based on a static generated dataset with JAMAR results to be able to share the user interface. The generated dataset version of the implementation is illustrated and explained in section 6.3.

4.5 Deploy: Release & Gather Feedback

During the deploy (7) step, the final developed visualization tool was provided to clinicians during an evaluation. The intended evaluation where clinicians use the tool in their clinics with patients could not be done in this study, due to time constraints. Instead, clinicians have been taught to use the tool by the visualization researcher and the tool was evaluated by letting the clinicians use their own computers, but without the patient present. The evaluation was done with the data of real patients of the evaluated clinicians. The technique used is the task-based analysis, where the participants are given tasks to perform with the tool. The evaluation tasks and a more detailed description of the clinician evaluation methodology can be found in section 7.1.

A qualitative approach was used to evaluate the efficacy of the newly developed tool. To answer the first research question RQ1, it is essential that the newly developed tool meets the needs of the clinicians that do the interactions with the visualizations. Thus, the observed clinicians are interviewed to learn how they experienced using the tool during their clinic. To avoid clinicians giving positive feedback because of experimental demand characteristics [44], the tool is mostly evaluated with clinicians that were not involved in any way during the development process.

The clinicians were asked to narrate their thoughts through the *thinking aloud* [45] protocol when they used the dashboard in their clinics. The results of this evaluation were analyzed through a *thematic analysis* [46] in Excel [47]. The results of the evaluation are reported in section 7.2.

The qualitative approach taken is not as robust and scientifically valid as a longer quantitative study at providing evidence for an improvement of standard of care provided because of the intervention with a visualization dashboard. A quantitative study of this sort does not fit the limited timeframe of this study, this limitation is further discussed in the discussion section. However, the qualitative approach taken still provides a clear first sign of the efficacy of using visualizations to visualize PROs and the results may function as a call to further evaluation in a future study.

This section has provided an overview of the design study methodology and subsequently, the 9-step framework that is adapted to conduct the design study. The methodology behind the precondition and core phases have been outlined. The next section will provide the results to the discover step.

5 Requirement Analysis

In this section, the requirements for the design are elicited based on the methodology outlined in section 4. First, a task abstraction is created in section 5.1 by analyzing the user observation to ensure that the design fits well with the workflow of the target user. Then, the clinician interview results are described in section 5.2. Section 5.3 presents an analysis of the patient interviews that give an insight into the requirements of patients with regards to the data visualization. While this research has pivoted away from focusing on the patient as a main target user as the study progressed, the patient is still expected to be exposed to the data visualizations during clinics, making the elicited patient requirements still valuable. Lastly, section 5.4 sums up the elicited requirements throughout this section in a table.

5.1 From User Observation to Task Abstraction

Two 4-hour clinic sessions were observed to understand how a clinician works and how new technology can fit the workflow of a clinician. During a single clinic session, roughly 10 patients are treated. During the second observed clinic, all actions undertaken by the clinician were logged and analyzed. Only the observations of this clinic were used in the analysis, resulting in 7 textual workflow descriptions of clinics with different patients. The workflows were combined into a task abstraction of an out-patient clinic process model illustrated in Figure 5.1. The model consists of 5 main tasks, designated by T1 to T5 that always occurred. Then during T2 and T4, subtasks were observed that occurred in some of the observations but not all. The number of observations made is indicated for each subtask, i.e., the JAMAR questionnaire (T2.1) was discussed in 2 out of 7 observations.

Each clinic starts with the greeting of the patient with a general question about how the patient is doing. The durations of these interactions differed strongly between patients and for that reason, no time indications are given in the task abstraction. After the greeting interaction with the patient, the medical summary of the patient is reviewed that is displayed on the screen as lines of text. Most of the medical history is written in a single text field where the clinician entered the digital letter received from the general practitioner of the patient. For 6 out of 7 observations, the clinician removed irrelevant data from this letter and proceeded to add new information that is gathered during the clinic with the patient. The clinician also validates the information given by the general practitioner with the patient to ensure that the data is still correct. In each observation, a pGALS physical functionality test (T3) was done to identify painful and swollen joints and limitations in functionality. This test goes through all the main joints throughout the body to test the range of motion available, with the goal of identifying pain and limits in physical functionality. After this test, the clinician would either continue the discussion of the medical summary (T2) or continue to discuss the next steps to take in treating the patient (T4). During this task, various aspects of discussion were observed. 2 out the 7 patients needed to be forwarded to the pharmacy for new medication, and 2 out of 7 patients needed to do lab tests that the clinician indicated through the EHR system. For the other patients, the clinician started a discussion on potential medication changes. Lastly, each clinic was concluded by greeting to the patient (T5).

The visualization dashboard will be most suited to assist the clinician during task T2.1, the discussion of the JAMAR questionnaire. By assisting the clinician during this task, the quality of the discussion will hopefully increase, which may also increase the number of occurrences of this task. The discussion of daily life (T2.2) and display of graphs (T2.4) could also be integrated into the dashboard solution. The questionnaire provides unique information about the quality of life of the patient, allowing for support during the discussion of daily life. Since the JAMAR questionnaire provides information on physical functioning, there is also an opportunity to assist clinicians during the pGALS test of functionality (T3). Furthermore, the dashboard could provide the clinician with information about medication prescriptions and adherence over time, allowing it to support the discussion of medication change (T4.2). To conclude, the task abstraction presented in Figure 5.1 illustrates the workflow of the clinician during a clinic and it illustrates how the dashboard has the potential to support the clinician during the JAMAR discussion, discussion of daily life, display of graphs and discussion of medication change. The next section delves deeper into the characteristics of the target user and their workflow through an interview.

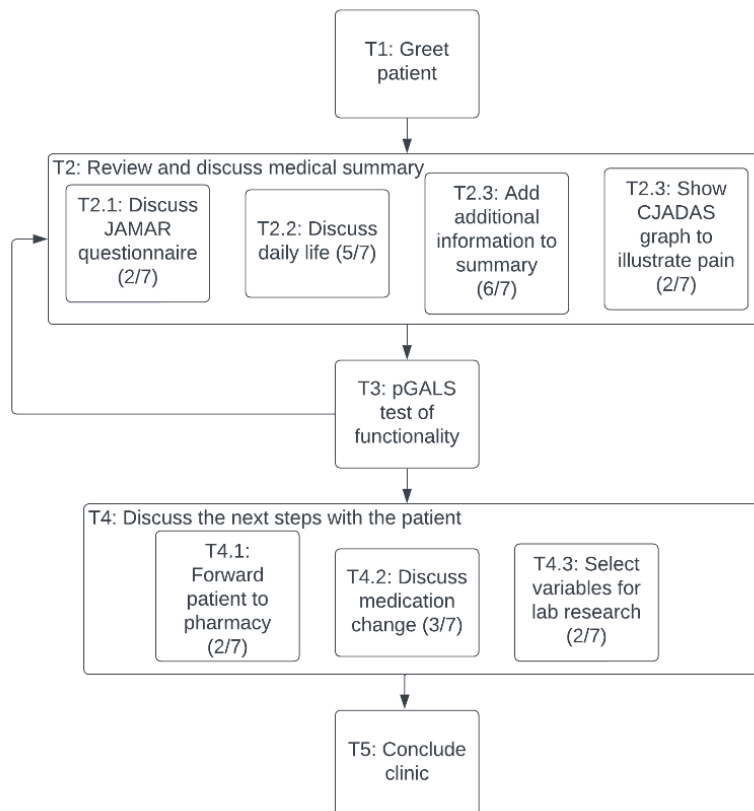


Figure 5.1 Task abstraction of out-patient clinic

5.2 Expert Requirements Interview

A pediatric rheumatologist is interviewed to get a better understanding of the domain and the requirements of the target user. This clinician is the same person that was observed in the previous section, allowing for interview questions aimed at clarifying observations made. The interview took place directly after the observation, allowing *critical incidents* to be pointed out that had been observed during the user observations. The first part of this section aims to clarify the user observation. Then, the current situation with the use of the PRO data is discussed and lastly, how this PRO data could be used more effectively is detailed. The results of this interview are used as requirements for the design in section 6.

5.2.1 Clarifying the User Observation: How Do pediatric rheumatologists work?

The clinician preferred to write and read stories while informing himself with medical information, as part of task T2, in the textboxes mentioned in the previous section. He mentioned that this is his style because he can write down and recognize the priorities of the complications better than a bullet point style. Later in the interview the clinician also mentions that this style of writing medical information allows him to automatically generate a letter to general practitioners, which he says some other clinicians struggle with. This letter is sent to the general practitioner of the patient whenever a new treatment plan is executed. This letter does not contain all considerations of treatments and diagnoses.

All pediatric rheumatologists have a similar setup where two computer monitors are used that are both visible to both the patient and the clinician. The monitor closest to the clinician is primarily used to work in the current EHR software called HiX, while the second monitor is primarily used to present data to the patient. The secondary monitor is also in closer proximity to the patients, which makes it more suitable for the purpose of presenting. The rationale behind this setup is to increase the transparency of the clinicians' actions regarding the computer system towards the patient. The clinician also mentioned that the patients can see almost all the data that was entered by the clinician on their own devices once they are logged into their healthcare portal. The only exempted data field is the 'diagnose differential' that contains the considerations and doubts regarding possible disease diagnoses.

The most important use of the HiX system is to integrate the learned information through the HiX system with the information known by the clinician and to present that information to the patient. With the presentation of information, the clinician would like to use the system to illustrate the progression over time, which is possible to a limited extent with the current system. To conclude, the clinician wants to keep using HiX. Therefore, the new solution should be usable alongside the existing EHR program instead of replacing it (**R1**).

5.2.2 *How is the JAMAR currently used?*

The clinician mentioned the JAMAR and the lack of its usage without the need of a question and mentioned that it is also a medical history check of some sorts, but that he does not use it currently. He estimated that the submission of the JAMAR by the patient requires around 15 minutes. Patients can enter the JAMAR on a device at home, on their mobile phone in the waiting room or on a computer of the hospital while waiting for their appointment. Occasionally the patients do not have enough time to complete the JAMAR while waiting for their appointment. In most of these cases the JAMAR is not completed after the clinic, as the patients have lost their motivation to complete the questionnaire while the decisions have already been made during the clinic. The clinician prepares all his clinics but does not use the results of the JAMAR in this preparation. To view the results of the JAMAR, several clicks are required by the clinician. The results can be observed without taking over the entire screen, which allows the clinician to keep on reading and adding medical information into the text boxes. This property of the current view of the JAMAR was spoken about favorably. However, the current view of the JAMAR results was described as unattractive with solely black on white text. There are no signals given on important values that require attention, thus forcing the clinician to look through all the answers for the important information. When asked about existing solutions to the current view of JAMAR, the clinician mentioned Qlik questionnaires that presented questions in a better way by changing the color of questions with strongly differentiating results. This Qlik questionnaire is done yearly to measure the quality of life of patients, in addition to the JAMAR questionnaires that are used. The clinician also mentioned 'kindtool' that uses grades of 1-10 that effectively signals the strongly differentiating values that require special attention. However, there are not yet any solutions to the JAMAR results presentation.

5.2.3 *How can the JAMAR be used more effectively?*

While the JAMAR was criticized by the clinician for its lack of inertia related questions, the potential data from the JAMAR questions was deemed as valuable. When asked for the most valuable questions, the clinician answered in a split-second that the last question of the JAMAR is the most valuable. This question 16 indicates whether a patient would be satisfied when the disease activity does not change in the upcoming months. The result of this question is central to other results of the JAMAR and can function as a call to action when a patient chooses 'no'. However, the clinician pointed out that the results of the JAMAR are still valuable when the patient chooses 'yes'. Then, the clinician mentioned that the Visual Analogue Scale (VAS) scores are important to clinicians while making decisions. This concerns questions 2, 6 & 15 of the JAMAR. Question 4 regarding the length of the morning stiffness was also deemed particularly valuable and for patients with more than 1 hour of morning stiffness, a JIA disease can never be regarded as inactive. The clinician could not choose any question that could be removed from the questionnaire, and thus all results of the JAMAR are deemed to be valuable for visualization (**R2**).

The clinician would like to see how the results of the JAMAR have changed between clinics for a single patient (**R3**). The data in the medical history textbox of the previous clinic is often

compared to the current condition of the patient while communicating with the patient during a clinic. Medical information about the condition that differs from the previous clinic can then be discussed with the patient. The clinician would like to be able to gain a similar oversight of the progression of JAMAR results over time so that answers that differ from previous measurements can be discussed with the patient. The clinician wants to see the delta of the extend of the value changes more easily (**R4**). Such a delta allows the clinician to see whether the condition of the patient is escalating or improving so that this can be discussed with the patient. During the interview, the clinician mentioned multiple times how the usage of different colors allowed him to identify differentiating values effectively, this principle can also be applied here.

The clinician would only use comparisons of the patients to the patient body and comparisons between patients for research purposes. The patients have little benefit in seeing such comparisons to other patients that are doing better or worse than them. The clinician is not interested in these comparisons during the clinics with patients. Here, the clinician put a priority on presenting patients with the JAMAR results over time instead of the comparison functions. This answer changes the scope of the project to be more focused on the visualization of JAMAR data to present it to patients. The use case of comparing patients during research timeslots will now fall outside of the scope. While a patient could be compared to patient body with JIA, comparisons to the healthy population are not possible because the healthy population does not submit JAMAR questionnaires. Currently, patients are primarily compared to other patients with the same disease diagnosis during research. The clinician showed an interest in finding out whether the patients would value such comparisons, such comparisons could potentially be investigated in a patient interview. It would also have to be investigated to what extend it is ethical to show patients how they compare to other patients. The clinician did not see any ethical issues with showing patients the progression of their JAMAR results over time, even when it clearly visualizes that the condition is escalating.

Discussing the JAMAR results with the patient was deemed as the most important usage of a potential new system. The clinician expressed an obligation to the patient to cover the results of the JAMAR during the clinic because the patients have to invest their time while do, they are annoyed with the questionnaire. By discussing the results with the patients more effectively, the reason for submitting the JAMAR will become more apparent to patients, motivating them during future submissions. The clinician would likely discuss the JAMAR with a new system at the end of the discussion of the medical history for roughly 2 minutes. By effectively using signaling outlying values, which was mentioned multiple times in the interview, the short usage timeframe of 2 minutes should be achieved (**R5**). The JAMAR results that are normal do not have to be discussed with the patients. However, with the deltas, the values that are progressing positively should also become apparent. The clinician was not interested in presenting the JAMAR data to any other actors than patients. The ability to annotate visualizations with text is not important for a new tool. The clinician would want to keep adding information to the existing textboxes in HiX on the primary monitor, while viewing the JAMAR visualizations on the secondary monitor during discussions of the JAMAR with patients, which corresponds with requirement R1.

Lastly, the clinician expressed an interest in seeing when the patient has visited the hospital and

for which dates there was a JAMAR submitted (**R6**). This would allow the clinician to discuss the adherence of the patient to the questionnaire submission. Besides the visualization of the JAMAR, the clinician would ideally like to see a timeline with medication changes on the X-Axis (**R7**) with varying data on the Y-Axis such as disease activity, pain or psychosocial wellness. This form of presentation is like the ‘kindtool’ patient chronicle that was mentioned earlier in the interview.

5.3 Patient Requirements Interviews

Three patients were interviewed together with their parents and one parent was interviewed without any patient input. Most of the answers were given by the parents of the patients, with the patients themselves confirming the statements of their parents and adding some of their experiences. Each parent-child pair is considered a ‘participant’. Some quotes given by the participants have been translated from Dutch to English.

All participants selected in this analysis had at least 2 years of experience with filling in the JAMAR questionnaire. All participants found the questionnaire clear and there were no confusing elements to the questionnaire. While narrating their experience with the questionnaire, two participants mentioned that the questionnaire is superfluous in some respects: “The questionnaire asks me whether I go to school for example, and then the doctor asks that question again during my clinic.” Another participant mentioned that some questions were not applicable to their child and that they would have to keep filling in ‘not applicable’. Two participants mentioned that they found the questions that are not directly related to their localized inflammation and pain were superfluous: “Some questions are a bit superfluous, when it is not about my wrist, then I am not that interested.”

When asked to narrate what they would like to learn and discuss with their doctors if the doctor were to have all data at the ready, three participants were able to formulate their answers after some time and one participant needed more specific questions to formulate an answer. One participant mentioned that they would like to see the effects of the medicine used on the results of their JAMAR and other measurements (**R7**). The participant was especially interested in the outcomes related to school in this report. They expressed a great desire to see the effects of the medications with their own eyes. Another participant was not interested in learning through visualizations at all: “I know how my child is doing, I don’t need graphs for that, as long as the doctor knows what is going on.” This participant mentioned that the doctor asks questions about topics that are already answered in the JAMAR questionnaire: “Often he asks about things like school, which I already answered in that JAMAR, so these questions are a bit double.” Two other participants shared the sentiment that it is more important for the doctor to understand their data well, than for them to be presented about graphs about patient well-being. However, they were still interested in learning how bad their disease is doing and if things are going in the right direction (**R3-R4**). None of the participants mentioned a desire to learn how their measurements compare to the patient population or healthy population in their answer to the open question. However, when asked more specifically about seeing how they compare to other patients or average healthy people, one out of the four participants expressed an interest. They

mentioned that they would like to know what the normal and healthy values are and whether they are moving towards that.

5.4 Elicited Requirements

Section 5 has presented the results of the user observation, clinician interview and patient interviews to come to a task abstraction and a list of requirements for the visualization system. The elicited requirements are summed up in Table 5.1. Section 6 will transform these requirements into a design.

No.	Requirement Description
R1	Used as an addition to their EHR program to learn information, not as a replacement
R2	The answers to all JAMAR questions need to be accessible in the dashboard
R3	Progression of PRO results over time
R4	Identify whether disease is improving or escalating
R5	Learn most important results within 2 minutes
R6	List the dates of clinics and whether a questionnaire was filled in for each date
R7	Allow for medication use to be compared to PRO results

Table 5.1 Requirements by number

6 Design

The elicited requirements have been transformed into a dashboard design as part of the ‘design’ step of the 9-step framework. This process started with sketching low-fidelity prototypes with a visualization focus group. Section 6.1 illustrates how initial visual encodings were created together with the visualization focus group. Section 6.2 describes how the findings from the visualization focus group and the elicited requirements were used to design a mockup design. This mockup design was then used to create a high-fidelity prototype that using programming interfaces, based on the mockup design. Section 6.3 describes the development of the final prototype that was implemented at a hospital as part of the ‘implement’ step of the 9-step framework.

6.1 Visualization Focus Group

A visualization focus group was held to come to a low-fidelity prototype that matches the needs of visualization users in the medical domain. The interviewed clinician from the previous section was present, as well as two medical researchers and 5 PhD students of medical studies. The focus group took place in an online environment, where the participants were exposed to a presentation of sketches. During the focus group, line graphs were widely seen as the best way to visualize temporal medical data. Bar charts are seen as inefficient, because of the need to see how medical data progresses and to see the direction of the data. Figure 6.1 illustrates an example of the bar chart and line chart that were shown to the focus group. Circular views were difficult to interpret for the focus group, indicating that out of the visualization methods identified in section 2.3, the stacking of line graphs is preferable over the use of circular views. The focus group was unanimous in the desire to see trends in measurements over time. The addition of emoticons to indicate the severity of measurements did not add anything for the focus group. While the literature review indicated that it is effective for patients at providing a more personal experience, it is not a good choice for medical experts. Thus, it became apparent that the requirements of data visualizations tailored for use by clinicians differ from requirements of data visualizations created for patients.



Figure 6.1 Low-fidelity sketches of bar-chart (left) and line-chart (right) of medical data

For color, traffic light colors are most used in the medical domain. Gradients with magenta were seen as less clear and alarming, such as the right graph illustrated in Figure 6.2. The red and green colors are uniformly understood as bad and good indicators by medical professionals. This corresponds with the findings in the literature review in section 2.2 that concluded that traffic light colors are useful for navigating the attention of clinicians. However, this feedback contradicts the finding in section 2.2 that described how gradations of magenta are also an effective color scheme to visualize symptom severity.

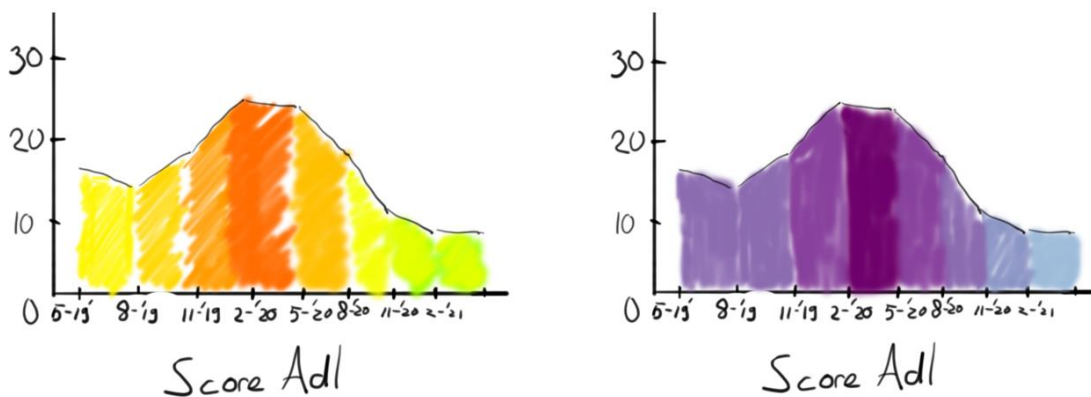


Figure 6.2 Sketches of area charts using stoplight color scheme (left) and magenta gradient color scheme (right)

A clinician suggested adding the average values of the healthy population to the visualizations, along with standard deviations from that healthy average. For clinicians with experience with statistics, this could give a clear indication of how good or bad a value is. The average and standard deviations could be added to line graphs by adding horizontal lines to the graph, so that each data point can be easily compared to the average in these graphs.

The Visual Analogue Scales that are part of the questionnaire indicate the severity of the disease, with a 10 being the most severe. The option was explored to visualize these scales as grades by inverting the number. Once again, this option may be effective for a patient dashboard, but it is seen as confusing for the clinicians that use the dashboard. Clinicians are used to seeing these scales and their accommodating scores in terms of severity and are focused on reducing these values.

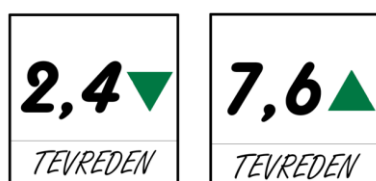


Figure 6.3 *Example visualization of satisfaction score as recorded through the VAS scale (left) or inverted into a grade (right)*

The focus group was eager to learn about medication usage while seeing the data visualizations. Thus, the inclusion of the medications used by patients on the timelines of visualizations was discussed. The JAMAR contains data about medication used medication adherence as well that can be used for this purpose. An existing method of visualizing medication usage on a timeline is the addition of horizontal lines below the timeline, where you use the X-axis that presents time. This way, the user can see during what periods medications were used while the graph above visualizes the results of the JAMAR questions

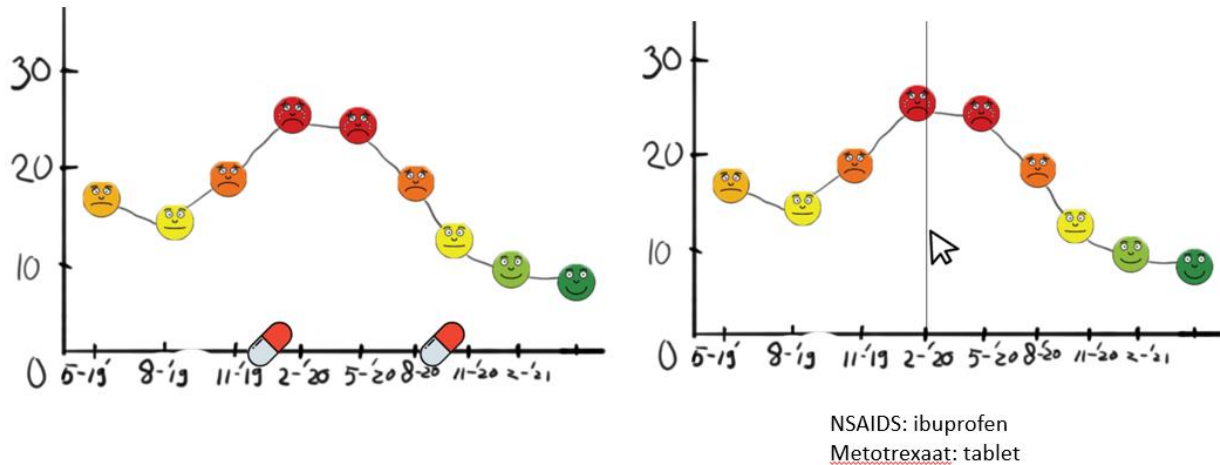


Figure 6.4 *Low fidelity sketches that illustrate medication changes (left) or medication used at the selected date (right)*

The findings of this focus group were used to create a higher fidelity mock-up design. The most important considerations are that line charts are the preferred method of visualizing temporal data, that the use of emoticons is not preferred, the stoplight color scheme is more effective than the magenta color scheme for indicating symptom severity and that VAS scores should be visualized just like they are recorded.

6.2 Mockup Design

Based on the derived requirements, a mockup design was created. The tool Adobe Photoshop was used to create an interface with modules that can be enabled and disabled quickly to allow for an evaluation session. An interface snapshot is illustrated in Figure 6.5. Section 6.2.1 will describe through which interaction techniques a clinician can select a patient to quickly receive a summary of the PRO data, which is described in section 6.2.2, and to then filter down to more detailed information, which is described in section 6.2.3, based on the summary. Lastly, section 6.2.4 goes over some of the rejected visualization techniques and explains the reason for the rejection.

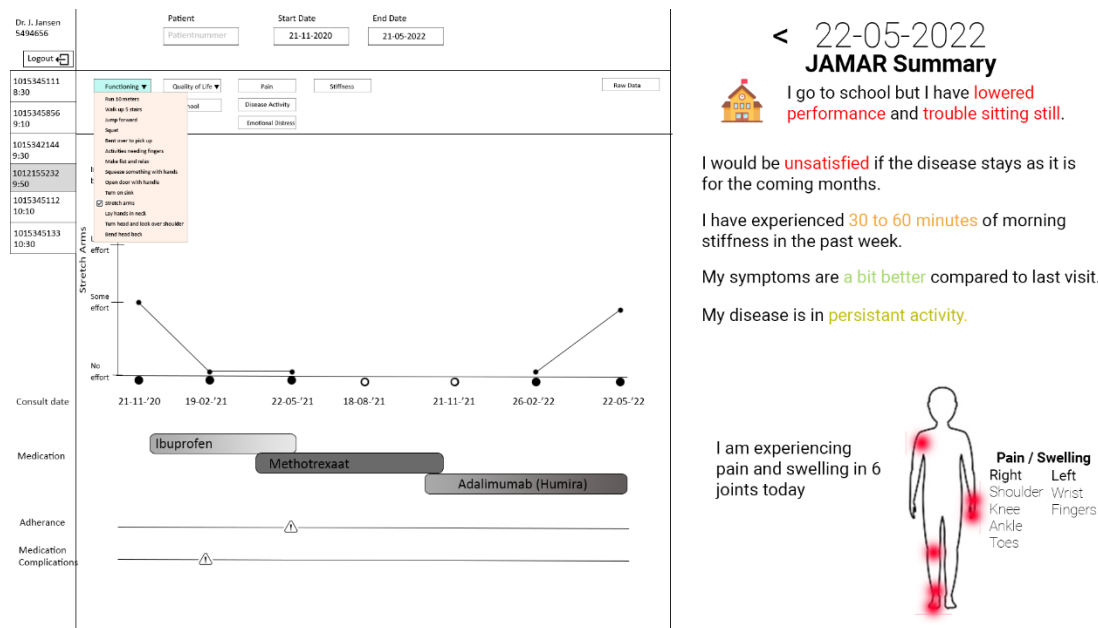


Figure 6.5 Mock-up design with narrative summary and one visualized attribute

6.2.1 Interaction techniques

Before describing the visualizations of the interface, the interaction techniques to acquire the PRO data will be detailed. Figure 6.5 illustrates how the user is logged in on the top-left of the screen. This is an important feature to protect the confidentiality of patient data. Below that, several patient numbers are displayed with the timestamps at which the clinician sees these patients in a clinic. This was done to accommodate for the clinicians' workflow where they see the patients in this sequential order, allowing them to quickly select the patients during the clinics to avoid an unnecessary loss of time. The patient number can also be manually entered in the patient number input field at the top of the screen. To acquire the patient number, the clinician would have to use their EHR software which is in correspondence with **R1**. Below that, the clinician can easily access different variables through buttons and drop-down menus for variables that consist of multiple sub-variables. Thus, these interaction techniques were purposefully designed in a way to allow the clinician to access the PRO data of all PRO questions (**R2**) as quickly as possible, based on the workflow of the clinician.

6.2.2 Narrative summary

The narrative was created on the right-hand side of the screen, denoted as the JAMAR summary in Figure 6.5. Because of the identified limited time available to use the dashboard, namely 2 minutes (**R5**), there was a need for a quick overview of the patient and the ability to narrow down to more detailed information. The purpose of this narrative is to give the user an initial

indication of how a patient is doing in the different aspects of the PRO questionnaire. For example, Figure 6.5 illustrates how the patient is suffering from morning stiffness and several painful joints, allowing the user to acquire additional information of these aspects with further use of the dashboard. This information is acquired in the right-hand side of the interface, the JAMAR summary. This summary should be the starting point for the clinician to get an indication of the most important values to investigate further. To achieve an obvious starting point, colors were used to denote parts that require attention. The summary in Figure 6.5 grabs the target users' attention by using the red color for the low school performance and dissatisfaction of the patient with the disease progression. So, by using the coloring technique in the narrative, the target user will quickly understand what attributes require further attention after which they can filter to achieve more detailed data about these attributes.

6.2.3 Timeline with line charts and medication data

The more detailed information is presented on the left-hand side of the screen which features a timeline with dates where clinics took place (**R6**). Empty dots above the clinic dates indicate missing data, i.e., the questionnaire was not filled in for this date. The timeline of medications prescribed to the patient is visualized below the timeline with bars (**R7**). Moments in time with bad medication adherence is visualized with the warning signs on horizontal lines in the bottom-left side of the screen. The interface illustrated in Figure 6.5 illustrates how the user has selected the 'Stretch arms' sub-variable from the 'functioning' supervariable. The selection results in the visualization of this attribute over time on the timeline (**R3-R4**) where this example illustrates that for some dates, this took some effort for the patient and other dates this took no effort.

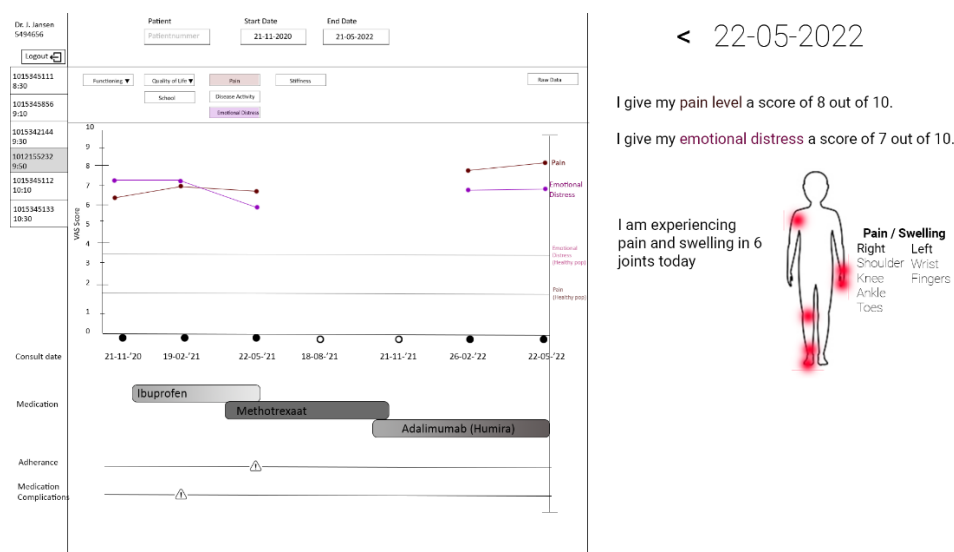


Figure 6.6 Mock-up design with pain and emotional distress selected

6.2.4 Rejected visualization techniques

Some visualization techniques such as the magnitude and green/red medication adherence dots were excluded based on received feedback. Figure 6.7 illustrates how three different attributes are visualized simultaneously over time through three magnitude charts. These magnitude charts display the VAS scores of the pain, disease activity and emotional distress variables on a 1 to 10 scale where large dots represent the high end of the scale and small dots represent the low end of the scale. This approach allows for the visualization of many attributes in a concise manner, but there it was excluded for its low accuracy. Furthermore, the medication lines using stoplight colors were excluded to prevent an overuse of color in the dashboard. The use of color is reserved for the most important attributes of the narrative and charts instead, to purposefully direct the user through the screen instead of scattering attention. Thus, the visualization techniques presented in Figure 6.7 were omitted in favor of the visualization techniques presented in Figure 6.5 and 6.6. The mockup design illustrated in these figures were used as the basis for the high-fidelity prototype that is created next.

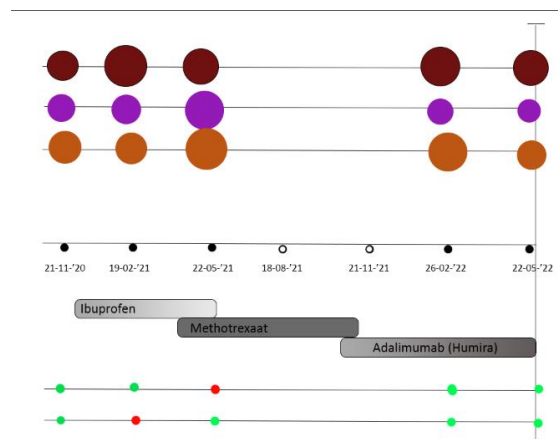


Figure 6.7 *Alternative visualizations of mock-up design that were rejected*

This section has introduced a mock-up design that adheres to the requirements elicited in section 5. The next section illustrates how this mock-up design has been implemented through a high-fidelity prototype.

6.3 Implemented high-fidelity prototype

This section describes the implementation of the high-fidelity prototype, PROVIs. First, section 6.3.1 describes the technologies used to create the prototype and context around the implementation. Section 6.3.2 describes the interaction techniques that enable the user to

navigate the dashboard. Next, section 6.3.3 describes how the narrative summary from the mock-up design has been implemented in the high-fidelity prototype. Section 6.3.4 describes how the timeline has been implemented through line charts. Lastly, section 6.3.5 goes over the differences between the mock-up design and the implemented prototype and finishes with a conclusion to the section.

6.3.1 Prototype context

The implemented prototype was developed based on the created mockup with the R shiny infrastructure. The existing infrastructure and experience with the R programming language made R the most logical choice of technology. The R shiny dashboard also makes use of web development technologies such as HTML and CSS, allowing for the customization of the interface. Furthermore, plot.ly visualizations were used to generate plots within the R shiny dashboard. Data was retrieved from an SQL server, through SQL statements. Tidyverse packages have been used for the data transformation.

The prototype is created to be used in addition to the existing EHR software that is used by the hospital. The prototype adds a layer of visualization to the PRO of the patient but does not replace annotating functionalities of the existing software. The prototype is designed to be used on a second screen alongside the EHR software that is opened on the main screen. The note taking of the clinician happens in their trusted EHR software. The rest of this section will describe the functionality of the prototype, along with the design choices behind the different aspects of the design.

6.3.2 Interaction techniques

The interaction techniques are largely adopted from the mockup design, with some alterations due to limitations. The top-left of the interface features a numeric input for the patient number with a ‘Go’ button that initializes a new patient. The user can copy this patient number from their existing EHR program (**R1**) or choose to manually enter this number. The selection of patients out of a list could not be achieved due to data access limitations. With this design, when the user hits the ‘Go’ button, all PRO results that correlate with the entered patient number are pulled from research data platform of the hospital. Furthermore, all dates are pulled where the patient had a clinic appointment at the hospital. With these clinic dates, a data frame is built where the PRO results are added to a clinic date if the PRO were submitted within one week of the clinic date. It is essential that the clinic dates are central to the visualization based on the elicited requirement **R6**. The clinic dates are listed in a drop-down menu above the narrative. Clinics where no PRO was submitted are visualized in red, giving the user a quick indication of all the existing clinics of a specific patient and for which clinics a PRO questionnaire was filled in.

Furthermore, a variable selection dropdown menu was added next to the patient number input to

allow the user to navigate through the various themes of the questionnaire (**R2**). Figure 6.8 illustrates the interface for which the ‘pain’ variable is selected for a fictive patient with patient number ‘10000001’. The clinic date ‘2018-07-01’ has been selected in the clinic date selection menu. These interaction techniques should allow the user to access the patient, desired attributes, and date in a timely manner without cluttering the interface with too much information.

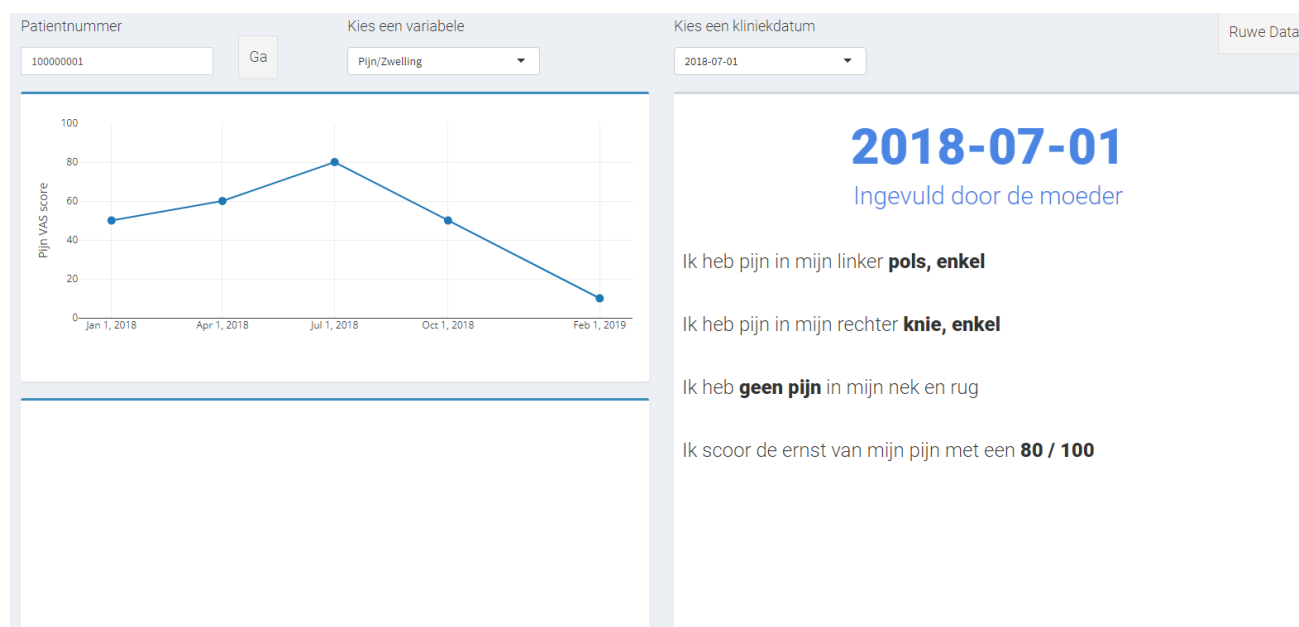


Figure 6.8 *Implemented user interface with ‘pain’ attribute selected*

6.3.3 Narrative summary

The narrative summary presents the PRO data by listing the most critical values first. First, the most important questions identified in the clinician requirements interview are listed first. E.g., whether the patient would be satisfied if the disease would not change is presented up top. These important questions can quickly inform the clinician about the most important attributes (**R5**), and then guide the clinician to use the attribute selection menu to zoom in on more detailed information for the related attributes.

6.3.4 Line charts with textual descriptions

The line chart at the top left visualizes the perceived level of experienced pain by the patient

over time (**R3-R4**). By selecting the clinic date, the user is presented with more detailed information for that specific date in the narrative on the right-hand side of the interface. The narrative tells the user that the patient is experiencing pain in their left wrist and ankle and right knee and ankle and that no pain in the neck or back was experienced. The use of stoplight color that was present in the mockup was omitted in the prototype to minimize the use of color. For the sake of completeness, a ‘Raw data’ button is included in the top-right of the interface. This allows the user to generate a table with each submitted PRO listed on a row, along with all relevant metadata.

To visualize the functionality scale consisting of 15 sub-questions and the quality-of-life scale consisting of 10 sub-questions, two different visualizations are used. Figure 3 illustrates how the top chart shows the progression of a cumulative functionality score that is based on the results of the 15 sub-questions. The user can then see how this cumulative score is built up by selecting sub-questions in the bottom visualization of Figure 6.9. The sub-questions can be selected by clicking on them in the legend of the graph. This approach allows for any number of the sub-questions to be visualized simultaneously, which makes use of the unique characteristic of clinicians of being able to consume large numbers of visualizations in a single screen.

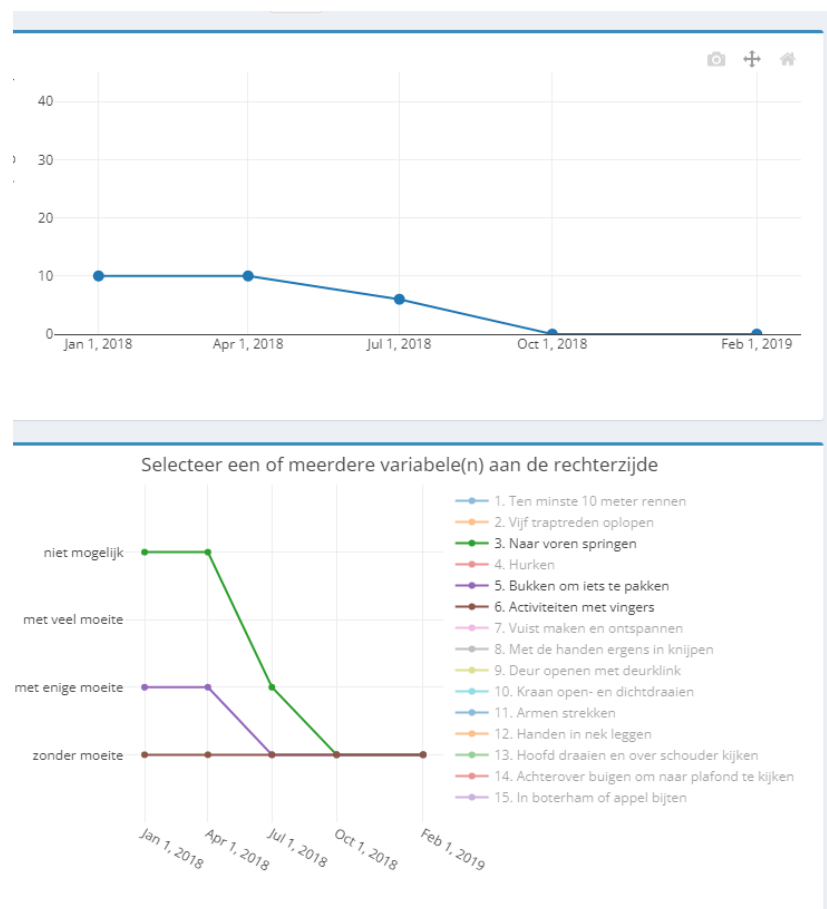


Figure 6.9 *Functionality visualization*

6.3.5 *Deviations from mock-up design*

The developed high-fidelity prototype lost some of the desired functionalities that were included in the mockup design due to time and data access constraints, the time restriction is further discussed section 8. The visualization of medication usage and adherence (**R7**) that was presented in the mock-up design through horizontal bars below the line charts could not be implemented due to technological constraints of R Shiny. Nevertheless, the implementation adheres to the elicited requirement by allowing this information to be accessed in the narrative. Furthermore, the visualization involving the human body with indications of pain through red circles was not included in the prototype due to time constraints. This design section has described how the initial sketches were transformed into a higher fidelity mockup design, that was used for the final developed prototype that is illustrated in Figure 6.8 and Figure 6.9. While some concessions have been made due to technological and time constraints, the implementation manages to adhere to the elicited requirements of section 5. The next section will describe the evaluation of this implementation.

7 Evaluation

This section explains the evaluation of the implementation created in section 6.3. First, section 7.1 describes the methodology of the evaluation process. Then, section 7.2 describes the results of the evaluation process through a thematic analysis. Lastly, section 7.3 gives an overview of the willingness to adopt the implementation by the clinicians.

7.1 Evaluation Methodology

The evaluation process is summed up in a process model in Figure 7.1. The figure illustrates the time expenditures of each evaluation task and is mostly linear in nature. The evaluation starts with a 15-minute tutorial of the dashboard, which is explained in section 7.1.1. Then the clinic preparation task is given to the clinician for two patient numbers, this task is described in section 7.1.2. The figure illustrates unless the participant indicated that they would not use the dashboard during their clinic, the ‘clinic with patient’ task is performed for two patients, this is described in section 7.2.3. Lastly, each section 7.2.4 explains the interview process that took place at the end of each evaluation.

The goal of the evaluation is to test the newly developed visualization tool for PRO results. The goal is explained to the clinician. The clinician is asked for permission to record the evaluation. Upon receiving approval, the audio is recorded through two separate devices. Then, the clinician is asked to start up their electronic health record program, just like they would in their clinic. Instructions are provided to access the dashboard that is running during the evaluation on the second monitor. Once the desired setup is achieved, the structure of the evaluation is explained.

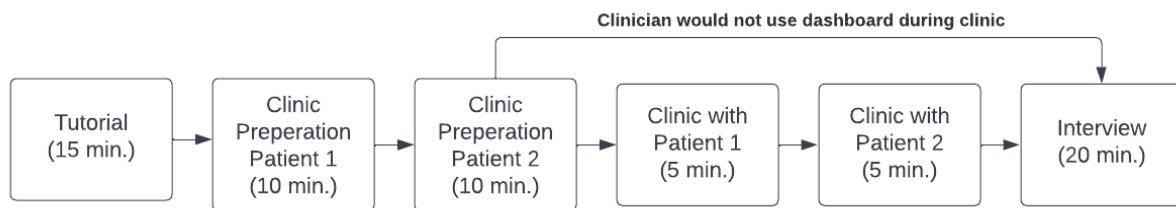


Figure 7.1 Evaluation steps with allotted time for each step

7.1.1 Dashboard tutorial

The evaluation starts with a tutorial of 10 minutes with the purpose of helping the clinician understand how the dashboard can be used. For this tutorial, the functionality of each user interface element is explained, starting with the patient number input field. The clinician is tasked to fill in a provided patient number to test this functionality. After the test patient is

generated, the functionality of the dropdown menu with clinic dates is explained. Then, the clinician is tasked to select the most recent clinic, for which a PRO questionnaire is filled in. After the correct clinic is selected, the information provided in the narrative is explained as a summary of the PRO results. Then, the clinician is introduced to the variable selection dropdown menu. Each of the selectable variables and the way that they relate to the questionnaire is shortly described. Then, the clinician is tasked to select the functionality scale variable. The generated graphs and narrative are explained, and the functionality of the secondary graph that two variables have is explained. Then, the clinician is tasked to select all variables by double clicking on any variable in the legend. The graph will illustrate that there was one peaking yellow line at a specific date, and the clinician is tasked to visualize only that line by double clicking on the yellow line in the legend. Then, the clinician is asked to include two more lines in the line graph by selecting the variables in the graph legend. After the graphs are understood, the clinician is tasked to select the pain variable in the variable selection. The dashboard will visualize the data related to pain and the clinician is asked whether they can spot the most recent date where the patient experienced pain. Upon successful identification of the date, they are asked to select this date in the clinic selection menu, to learn in what joints the patient experienced pain at that date. Lastly the clinician is introduced to the raw data button to view the answers of the selected date in tabular format. The clinician is given the opportunity to ask any questions about functionalities of the dashboard that are unclear.

7.1.2 *Clinic preparation task*

The evaluation itself takes place in two stages and the clinician is asked to narrate their thoughts in their own language through the *thinking aloud* protocol during the clinic. The first stage commences with a story about how the clinician is preparing for a clinic with a patient that the clinician will be seeing soon. The clinician is told that they are about to see a specific patient and that they are taking some time in the preparation of the clinic to study the results of the PRO questionnaire. This patient is a patient that the clinician will see in one of their next clinics in their real-world practice. The name of the patient is given to the clinician, and it is up to the clinician to copy the patient number from the electronic health record program to the patient number input of the dashboard. Entering a patient number will generate the relevant data for the patient so that the clinician can start their preparation. The clinician is told that to prepare for the clinic, they would like to learn how the patient is currently doing in their day-to-day life and what symptoms the patient is currently experiencing. Furthermore, the clinician would like to know if these aspects have been becoming better or worse over time.

7.1.3 *In-clinic usage task*

During the second stage of the evaluation, the clinician is told to imagine that they are currently in a clinic with the patient that they prepared for. During this clinic, the clinician would like to allocate a few minutes to going over the results of the PRO with the patient. The clinician will go over the most recent PRO, pointing out parts of the patient reported outcomes that may be

worth discussing with the patient. If the clinician does not generate any line-graphs by themselves, they are told that they are curious about the progression of the disease over time and to discuss this progression with the patient. Lastly, the clinician is interested in discussing to what extent the PRO questionnaire was filled in, and the effect the level of completeness of the PRO data has on the generated dashboard.

7.1.4 Follow-up interview

After the experiment, the following questions are asked to the clinician. Observations that were made during the experiment, are pointed out to the clinician, following the *critical incident* technique used in the requirements interview. The questions are asked in Dutch. The first question Q1 aims to acquire general thoughts about the prototype. Follow-up questions were asked to get detailed responses in cases where this was necessary. The follow-up questions can be found in Appendix D. Questions Q2 & Q3 serve to find strong and weak points of the prototype by pointing to critical incident observations that were made during the evaluation. Lastly the participants are asked to what extent they are willing to adopt the dashboard in their daily practice, to learn about the technology adoption.

***Q1:** What did you think about using the dashboard to learn about the patient reported outcomes?*

***Q2:** A specific incident is pointed out where the dashboard seemed to provide a valuable addition to understanding the patient reported outcomes, and the clinician is asked to what extent he felt supported by the tool at that moment.*

***Q3:** A specific incident is pointed out where clinician experienced an issue with the dashboard, and the clinician is asked whether he felt sufficiently supported by the tool at that moment.*

***Q4:** Would you use the revised version of this dashboard in your next clinic?*

***Q4.1:** During the preparation of the clinic, or during the clinic itself?*

The results of the evaluation are analyzed through a *thematic analysis* [46] where data extracts are linked to codes in section 7.2.2. These codes are matched to identified themes in section 7.2.3, which will be presented in a thematic map. The thematic analysis is done in excel using established steps and methodology [47]. The thematic analysis is exploratory, meaning that the themes are still unknown and need to be discovered from the data. Section 7.2 details the results of this thematic analysis by first listing all codings in bold, with some quotes as evidence.

7.2 Thematic Analysis Results

This section presents the results of the thematic analysis of the evaluation with the 5 participants. Section 7.2.1 gives a short description of the sample size. Section 7.2.2 lists the codings that have been identified from the transcriptions of the evaluations and interviews. Lastly, section 7.2.3 presents how these codings have been mapped to three identified themes.

7.2.1 Sample Size

Evaluations were done with four pediatric rheumatologists and one pediatric rheumatology nurse that works with patients with JIA, for a total of five participants. Out of the four pediatric rheumatologists, one was exposed to the project during the requirements elicitation phase. The remaining three pediatric rheumatologists were aware of the project but were not exposed to any detailed information about the project before the evaluation. All participants see patients, that fill in JAMAR questionnaires, on a weekly basis. Thus, for each participant, two patients could be selected that are familiar with the JAMAR questionnaire.

7.2.2 Codings

Each coding found is listed in bold, with some example quotes to provide proof. Codes are only included if two or more participants mentioned something related to the coding.

Missing comparison to objective measure

Three participants mentioned a missing comparison to objective measure. “So you don't have a cJADAS, which is a bit of a shame. That way I keep having to look back and forth to HiX. That's what we always do, compare the scores of the patients with the cJADAS.” The cJADAS refers the assessment of disease activity of the patient. Participant 4 also mentioned a desire to see this score added in a graph in the same view. During the evaluation, participant 1 mentioned that the patient that was visualized has a large disconnect between their subjective experience and the measured active joints. Participant 1 immediately wanted to mention that the dashboard displays the subjective score, but that the clinician would love to compare the subjective scores

of the patient with the more objective measurements that exist. This sentiment was repeated by participant 3 who mentioned: Participant 2 was the only one wanted to see lab results in the same dashboard, to compare these to the PRO.

Ease of understanding

Three participants mentioned something regarding the ease of understanding of the dashboard. Participant 1 even took the initiative to go over all the functions of the dashboard without the need of any instructions. All participants needed less than 10 minutes to understand the functions of the dashboard. Participant 5 repeated that the dashboard has a “very clear overview in a fast way”, three times. Participant 3 mentioned that the “The program is very intuitive, and it is easy to find things.”

Information cluttering

“It's a lot in one graph to be honest. For her it is very important how she feels, the line graph shows how she felt very depressed at a time and that it is going better. So there is a lot of extra info, let's say..”

Long loading times

Three participants were slightly annoyed by the initial loading time of the prototype. For participant 1, the booting process took longer than expected; roughly 1 minute. During this time, the participant started on a bit of work on the EHR program opened on their other screen. The loading of the first patient took up to 15 seconds for some participants.

Quick overview through narrative

Four participants mentioned how a general overview can be acquired quickly through the dashboard. “I have a lot of new patients that I take over from my colleague, and this allows me to quickly get an overview of how the patient is doing.” None of the participants had negative feedback about the quantity of information that was presented in the summary. Thus, the target user seems to be comfortable with consuming large amounts of information in a single view.

Initializing patients

All participants were able to successfully insert the patient number that was provided in the tutorial in the patient number input box. 3 participants pressed the ‘Enter’ key to attempt to load

the patient after filling in the number. Thus, to make the interface more intuitive, the dashboard should allow for the generation of a patient by pressing ‘Enter’, besides the ‘Go’ button.

Clinic dates to see questionnaire adherence

All participants found the view of clinic dates to be an effective representation to see recent clinics and whether a questionnaire was filled in. Participant 1 immediately saw that for the last 3 clinics, no JAMAR questionnaire was filled in, which is common with patients suffering from depression. “Oh, you can see here that she did not fill in the questionnaire for the last 3 clinics, we see that a lot with patients dealing with depression. But that is also useful input, you know. Then I can discuss this with the patient when I see her next week.” Participant 2 found the overview of dates effective but would have liked to also see longer intakes and daycare to create a more complete overview of the patient journey. This sentiment was not mentioned by any other participants.

Missing detail in narrative

Participant 4 could not identify a specific patient satisfaction question of the questionnaire that was at the bottom of the summary. The information about medication usage and adherence to medication that was visualized in the narrative summary, could also not be found by participants 1, 2 and 4.

Clear overview of chronology

The main line charts of the main variables were perceived as clear and intuitive by all participants. Participant 3 did raise a question about the scaling of the dates of the x-axis, which could quickly be clarified. The visualized timeline was immediately clear to all other participants and perceived positively. Participant 2 mentioned that “You get a lot better view of the chronology of the patient, and it shows data about things you don't often discuss with the patient with regards to emotion.”

Ability to match up graph peaks

Two participants mentioned how they felt enabled in matching up graph peaks. Participant 3 was able to correlate the specific visualization of depressive mood sub questions with other measurements. This participant would also like to see a visualization of all inflamed joints over time.

Unintuitive multivariate line graph interaction

The interaction functionality of the multivariate chart did not feel intuitive for 3 participants. Participants 2 and 4 were not able to successfully visualize all lines by double clicking on the legend. Participant 5 mentioned that it was “a bit of a shame that you have to click through all the functionalities to see how they progressed”. Adding an indication of functionalities of interest to the narrative may reduce the number of clicks for the user to acquire this information.

Unable to identify medication usage and adherence

The information about medication usage and adherence to medication that was visualized in the narrative summary, could also not be found by participants 1, 2 and 4. Participant 2 mentioned that “The medication adherence is very important, and it does not come to my attention in the narrative enough.” Participant 5 would have liked to have seen a greater amount of detail in the narrative that supported the functionality scores of the patient. The participant wanted to see the sub questions where a patient does not do well highlighted in the narrative, even if this increases the amount of information on the screen.

Pressing ‘Enter’ not working

Participants pressed the ‘Enter’ key to attempt to load the patient after filling in the number. Thus, to make the interface more intuitive, the dashboard should allow for the generation of a patient by pressing ‘Enter’, besides the ‘Go’ button.

7.2.3 Thematic Mapping

The codes identified in the previous section have been linked to three main themes and this mapping has been illustrated in a thematic map. Figure 7.1 illustrates how the main themes are strong aspects, weak aspects, and missing functions. It becomes clear that the dashboard prototype succeeds at providing a clear overview of the chronology of attributes, is easy to understand, and that the narrative provides a quick overview of the results.

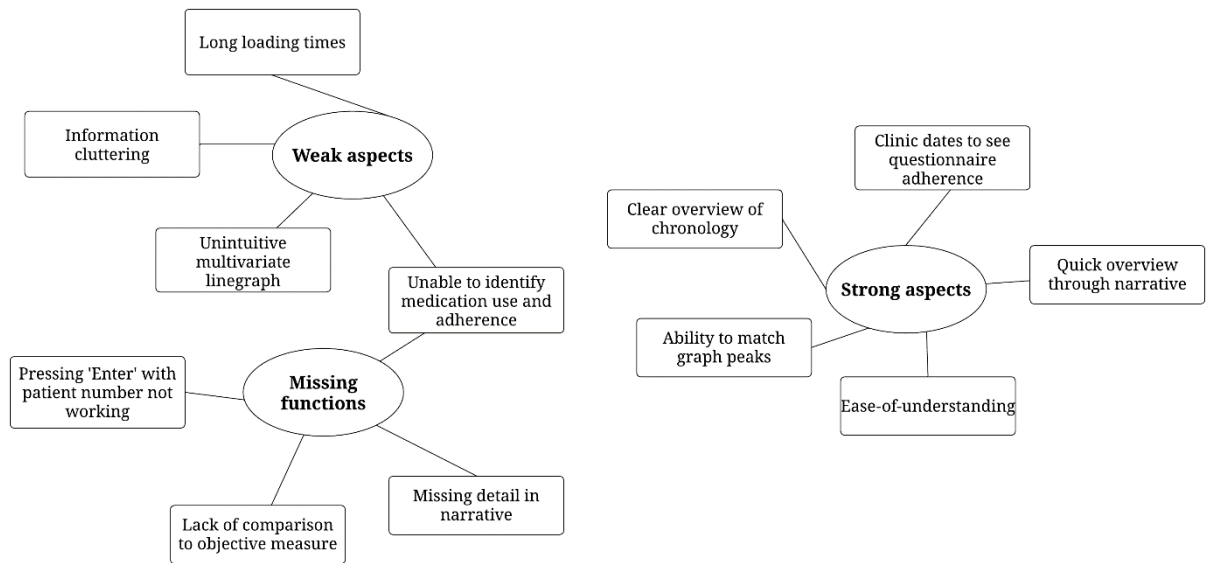


Figure 7.1: Thematic map

7.3 Technology adoption

The participants have been inquired into their willingness to adopt the technology in their practice. All 5 participants stated that they would use the dashboard for the preparation of their clinics. Participant 1 was worried about reduced ease of use while using the dashboard at home because of a setup with one screen but expressed a strong desire to use the dashboard. Participant 3 explained how they see many new patients and how the dashboard is especially helpful in that case: “I have a lot of patients that are new to me, and I lose a lot of time to get a first impression of how the patient is doing. This dashboard would save me so much time.”

Participants 4 and 5 stated that they would use the dashboard during the clinics with patients as well. Participant 4 mentioned that “I would use the graphs in the same way as I just did. I think people like to see graphs.” Participant 3 mentioned that they would have the dashboard open by default but would occasionally point out relevant visualizations to patients that were identified in the preparation of the clinic. Participants 1 and 2 stated that they would likely not use the dashboard in the clinics with patients, because they do not like to use digital tools while communicating with patients at all. They made it clear that this was not because of the tool, but because of the style of communication of the clinician. Participant 1 mentioned that “It has nothing to do with the functionality of the dashboard, but it is just my style to not have a computer in front of me. What I do is that I grab a piece of paper and draw two lines to show: this is how you think your disease is developing and this is how I think your disease is developing.” Table 7.1 provides an overview of the willingness to adopt the dashboard by all

participants.

While writing this thesis, sometime after the evaluation phase, the connectors of the hospital mentioned that the dashboard is now live and accessible to the team of pediatric rheumatologists. The implementation is being tested for 3 to 6 months, after which additional functionalities will be added.

<i>Participant</i>	<i>P1</i>	<i>P2</i>	<i>P3</i>	<i>P4</i>	<i>P5</i>
<i>Use in clinic preparation</i>	Yes	Yes	Yes	Yes	Yes
<i>Use in clinic</i>	No	Yes	No	No	Yes

Table 7.1: *Technology adoption for clinic preparation and in-clinic use*

8 Discussion

First, the limitations of this study are discussed. Then, the most important lessons that can be learned from this study are presented. The section finishes with a conclusion of this thesis.

8.1 Limitations

A limitation of this research was that the user observation was done with the researcher in view, as this could not be avoided with the confidential nature of patient clinics. Some actions taken on the screen by the clinician have been missed and not properly recorded because the screen could not be recorded due to confidentiality and the researcher had to sit at a distance to achieve fly-on-the-wall. The workflow of a clinician during a clinic with a patient was analyzed, and the dashboard was developed to meet this workflow. However, the evaluation illustrates how the use-case of clinic preparation without a patient present will see a higher adoption of the technology. Thus, future research could analyze the clinic preparation process and design a dashboard that fits this workflow better.

Another limitation is that of limited generalizability of the results. The evaluation illustrated how the participants have different preferences and use-cases for the dashboard. Since the department for which this dashboard was created consisted of only 5 clinicians, the number of participants that could be used for requirements interviews and the evaluation was limited. Furthermore, one clinician was inquired during the requirements interview. Considering how heterogeneous the group of clinicians was, it is safe to assume that the requirements of the interviewed clinician did not reflect the requirements of all pediatric rheumatologists. Increasing the number of participants for both the requirements interview and the evaluation would sketch a more accurate representation of the target user. Future studies could target PROs in different hospitals and with different diseases than JIA, to investigate whether the findings are similar.

The evaluation made it clear that the visualization and interaction design of multiple lines in a single line chart was not successful at visualizing the progression of 10 to 15 questions over time. Future iterations of the dashboard could attempt to use circular views such as a spider web visualization to tackle this problem. Furthermore, the evaluation indicated that additional detail in the narrative should be able to alleviate this problem. The evaluation also made it clear that future iterations of the prototype should include the subjective scores by the clinician so that these can be compared to the PRO. This could be achieved by adding an additional line-chart to the existing line-charts with PRO scores. Furthermore, the addition of lab results and medication data should be explored.

For this study, planned evaluations with patients were not feasible due to the restricted time available. Therefore, the clinicians could not be observed while using the visualizations in their real-world clinics with patients, despite them being a part of the initially intended target user group. The developed prototype is missing features that were part of the mockup design such as the medication timelines and puppet visualization.

Confidentiality of the patients is essential for the reporting of this project. Therefore, the interface of the final prototype with real patient data cannot be shown in the design section. Instead, an earlier version of the prototype is illustrated with mock patient data that is generated in a way to attempt to represent a potential patient.

8.2 Lessons learned

Based on the results of this study, there are some lessons that should be considered for future design studies with medical data and especially for future design studies with PRO data specifically. First, it is important for clinicians to be able to compare the subjective experience that is reported by the patient in a PRO with their own subjective score. Second, clinicians have limited time to use dashboards in their daily practice and an effective dashboard design should therefore provide the most important attributes as quickly as possible. This study showed how using a narrative summary received positive reactions when it comes to learning the most important attributes immediately. Third, it is vital to establish a feasible timeline early on that is agreed upon with all collaborators so that all research goals can be achieved. This study could no longer evaluate the perception of patients due to time constraints. A more feasible timeline would have prevented this. Lastly, while clinicians can consume many simultaneous visualizations effectively, stacking line charts seemed too cluttering to convey the progression of multivariate constructs. Future studies could look at an effective way to visualize large number of variables in a concise image for clinicians.

8.3 Conclusion

This study has provided a clear insight into the problem with PROs; clinicians find existing presentations of PRO results involving long lists of textual sentences difficult to consume, causing them to ask the questions directly to the patient instead. Consequently, patients find the PRO reporting to be a time-wasting activity, reducing their adherence to PRO questionnaires. To tackle this problem, a design study was performed at the pediatric rheumatology department of a Dutch hospital by following a 9-step framework. The literature review identified visualization techniques that can be used to visualize for clinicians, such as the vertical stacking of horizontal timelines and traffic light colors. Furthermore, the current state of PRO visualization has been researched and no studies could be identified that visualize PRO data for clinicians. This study has provided a clear insight into the needs and practices of pediatric rheumatologists with regards to PROs. These clinicians can consume large quantities of data in a single interface and desire to learn PRO results in a 2-minute timeframe so that they can direct their focus to their patients. A method of PRO visualization (PROVis) was presented that focused on the clinic dates that allows clinicians to see to what extent PRO questionnaires have been filled in, facilitating questionnaire adherence conversations with their patients. PROVis combines a narrative summary with line graphs with an adaptive narrative description centered around clinic dates. Evaluations have illustrated how the visualizations are received positively by the domain experts. The dashboard could be even more effective by adding in measures by

clinicians for comparison to the PROs. Furthermore, a more effective visualization technique that can visualize 15 categorical attributes over time would improve the intuitiveness of PROVis. Future research could apply the method of visualization to a different domain with different PROs, to validate the effectiveness of the visualization techniques with a different type of disease. Future work could also validate the results through quantitative evaluation by asking a large group of patients about their levels of satisfaction with the PRO process; before and after intervention with PROVis.

8.4 Ethical Considerations

This thesis has studied human subjects during the user observation, clinician requirements interview, patient requirements interview and the final evaluation. Ethical approval for these human experiments has been regulated by the main researcher, the daily supervisors at the hospital and the ethical committee of the hospital. Before the execution of the research methodology, the daily supervisors at the hospital reached out to the ethical committee of the hospital to approve of the research methodology. The ethical committee concluded that the research methodology did not require an official approval by the MREC NedMec review committee. The reasons provided were the goal of improving healthcare and the marginal need for behavior change or time investment from the human subjects.

The research methodology did not gather any personal information of human subjects to comply with GDPR and keep the gathered data anonymous. Moreover, interview recordings have remained within the hospital environment through a virtual machine that was provided by the hospital. All participants have provided consent to the researcher before the interviews and evaluations. For patient interviews, consent from the clinician that was responsible for the patient was also acquired.

References

1. Fayers, P.M., Machin, D.: *Quality of Life: The assessment, analysis and reporting of patient-reported outcomes*. John Wiley Sons Inc. (2016).
2. Deshpande, P., Sudeepthi, B.L., Rajan, S., Abdul Nazir, C.: Patient-reported outcomes: A new era in clinical research. *Perspect. Clin. Res.* (2011).
3. FDA: Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. *Fed Regist.* 74, 65132–65133 (2009).
4. Mercieca-Bebber, R., King, M.T., Calvert, M.J., Stockler, M.R., Friedlander, M.: The importance of patient-reported outcomes in clinical trials and strategies for future optimization. *Patient Relat. Outcome Meas.* (2018).
5. Munzner, T.: *Visualization Analysis & Design*. Taylor & Francis Group, LLC (2015).
6. Ware, C.: *Information Visualization - Perception for Design*. (2008).
7. Aigner, W., Miksch, S., Schumann, H., Tominski, C.: *Visualization of Time-Oriented Data*. (2011).
8. Sedlmair, M., Meyer, M., Munzner, T.: Design study methodology: Reflections from the trenches and the stacks. *IEEE Trans. Vis. Comput. Graph.* (2012).
9. De Croon, R., Klerkx, J., Duval, E.: Design and evaluation of an interactive proof-of-concept dashboard for general practitioners. In: *Proceedings - 2015 IEEE International Conference on Healthcare Informatics, ICHI 2015* (2015).
10. West, V.L., Borland, D., Hammond, W.E.: Innovative information visualization of electronic health record data: A systematic review. *J. Am. Med. Informatics Assoc.* (2015).
11. Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., Shneiderman, B.: LifeLines: using visualization to enhance navigation and analysis of patient records. *Proc. AMIA Symp.* (1998).
12. Barsh, A.: *The Big Book of Dashboards: Visualizing Your Data Using Real-World Business Scenarios*. Ref. User Serv. Q. (2017).
13. Wang, T.D., Plaisant, C., Quinn, A.J., Stanchak, R., Shneiderman, B., Murphy, S.: Aligning temporal data by sentinel events: Discovering patterns in Electronic Health Records. In: *Conference on Human Factors in Computing Systems - Proceedings* (2008).
14. Klimov, D., Shahar, Y., Taieb-Maimon, M.: Intelligent selection and retrieval of multiple time-oriented records. *J. Intell. Inf. Syst.* (2010).
15. Shahar, Y., Cheng, C.: Intelligent visualization and exploration of time-oriented clinical data. *Top. Health Inf. Manage.* (1999).
16. Shahar, Y., Goren-Bar, D., Boaz, D., Tahan, G.: Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions. *Artif. Intell. Med.* (2006).
17. Dowding, D., Randell, R., Gardner, P., Fitzpatrick, G., Dykes, P., Favela, J., Hamer, S., Whitewood-Moores, Z., Hardiker, N., Borycki, E., Currie, L.: Dashboards for improving patient care: Review of the literature, (2015).

18. Gschwandtner, T., Aigner, W., Kaiser, K., Miksch, S., Seyfang, A.: CareCruiser: Exploring and visualizing plans, events, and effects interactively. In: IEEE Pacific Visualization Symposium 2011, PacificVis 2011 - Proceedings (2011).
19. Kosara, R., Miksch, S.: Visualization methods for data analysis and planning in medical applications. In: International Journal of Medical Informatics (2002).
20. Aigner, W., Miksch, S., Müller, W., Schumann, H., Tominski, C.: Visualizing time-oriented data-A systematic view. Comput. Graph. (2007).
21. Buono, P., Aris, A., Plaisant, C., Khella, A., Shneiderman, B.: Interactive pattern search in time series. In: Visualization and Data Analysis 2005 (2005).
22. Zhang, Y., Chanana, K., Dunne, C.: IDMVis: Temporal Event Sequence Visualization for Type 1 Diabetes Treatment Decision Support. IEEE Trans. Vis. Comput. Graph. (2019).
23. Keim, D.A., Schneidewind, J., Sips, M.: CircleView - A new approach for visualizing time-related multidimensional data sets. In: Proceedings of the Workshop on Advanced Visual Interfaces AVI (2004).
24. Tominski, C., Schumann, H.: Enhanced interactive spiral display. In: Proceedings of The Annual SIGRAD Conference (2008).
25. Consolaro, A., Morgan, E.M., Giancane, G., Rosina, S., Lanni, S., Ravelli, A.: Information technology in paediatric rheumatology. Clin. Exp. Rheumatol. (2016).
26. Reeve, B.B., Hays, R.D., Bjorner, J.B., Cook, K.F., Crane, P.K., Teresi, J.A., Thissen, D., Revicki, D.A., Weiss, D.J., Hambleton, R.K., Liu, H., Gershon, R., Reise, S.P., Lai, J.S., Cella, D.: Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Med. Care. (2007).
27. Harrison, L., Skau, D., Franconeri, S., Lu, A., Chang, R.: Influencing visual judgment through affective priming. In: Conference on Human Factors in Computing Systems - Proceedings (2013).
28. Hakone, A., Harrison, L., Ottley, A., Winters, N., Gutheil, C., Han, P.K.J., Chang, R.: PROACT: Iterative Design of a Patient-Centered Visualization for Effective Prostate Cancer Health Risk Communication. IEEE Trans. Vis. Comput. Graph. (2017).
29. Stonbraker, S., Porras, T., Schnall, R.: Patient preferences for visualization of longitudinal patient-reported outcomes data. J. Am. Med. Informatics Assoc. (2020).
30. A., P., R., C., N., F., M., N., D., N., B., O., D., P., J., S., A., I., M., S.: The PROBE Data Dashboard - A new tool for real-time data visualization and analysis of patient reported outcomes. Haemophilia. (2018).
31. Elshehaly, M., Randell, R., Brehmer, M., McVey, L., Alvarado, N., Gale, C.P., Ruddle, R.A.: QualDash: Adaptable generation of visualisation dashboards for healthcare quality improvement. IEEE Trans. Vis. Comput. Graph. (2021).
32. Li, R., Hoque, E., Carenini, G., Lester, R., Chau, R.: ConVIScope: Visual Analytics for Exploring Patient Conversations. In: Proceedings - 2021 IEEE Visualization Conference - Short Papers, VIS 2021 (2021).
33. Prakken, B., Albani, S., Martini, A.: Juvenile idiopathic arthritis, (2011).
34. Thierry, S., Fautrel, B., Lemelle, I., Guillemin, F.: Prevalence and incidence of juvenile idiopathic arthritis: A systematic review, (2014).

35. Ravelli, A., Martini, A.: Juvenile idiopathic arthritis, (2007).
36. Filocamo, G., Consolaro, A., Schiappapietra, B., Dalprà, S., Lattanzi, B., Magni-Manzoni, S., Ruperto, N., Pistorio, A., Pederzoli, S., Civino, A., Guseinova, D., Masala, E., Viola, S., Martini, A., Ravelli, A.: A new approach to clinical care of juvenile idiopathic arthritis: The juvenile arthritis multidimensional assessment report. *J. Rheumatol.* (2011).
37. Wulffraat, N., Kamphuis, S., Swart, J.F., Vastert, S., Van Dijkhuizen, P., van Pelt, P., van Dijk-Hummelman, A., Consolaro, A., Bovis, F., Ruperto, N.: The Dutch version of the Juvenile Arthritis Multidimensional Assessment Report (JAMAR). *Rheumatol. Int.* (2018).
38. Munzner, T.: A nested model for visualization design and validation. In: *IEEE Transactions on Visualization and Computer Graphics* (2009).
39. Blomberg, J., Giacomi, J., Mosher, A., Swenton-Wall, P.: Ethnographic field methods and their relation to design. *Participatory design: Principles and practices. Particip. Des. Princ. Pract.* (1993).
40. Longhurst, R.: Semi-structured interviews and focus groups. *Key methods Geogr.* 3, 143–156 (2003).
41. Crandall, B., Hoffman, R.R., Klein, G.: *Working Minds: A Practitioner's Guide to Cognitive Task Analysis.* MIT Press (2006).
42. Salmon, P., Jenkins, D., Stanton, N., Walker, G.: Hierarchical task analysis vs. cognitive work analysis: Comparison of theory, methodology and contribution to system design. *Theor. Issues Ergon. Sci.* (2010).
43. Ambler, S.: *Agile Modeling: Effective Practices for eXtreme Programming and the Unified Process.* (2002).
44. Intons-Peterson, M.J.: Imagery paradigms: How vulnerable are they to experimenters' expectations? *J. Exp. Psychol. Hum. Percept. Perform.* (1983).
45. Nielsen, J.: Usability engineering. In: *Computer Science Handbook, Second Edition* (2004).
46. Braun, V., Clarke, V.: Using thematic analysis in psychology. *Qual. Res. Psychol.* (2006).
47. Ronan, B., Gallagher, G.: Using Microsoft Excel to code and thematically analyse qualitative data: a simple, cost-effective approach. *ALL Irel. J. High. Educ.* (2016).

Appendix A. Data Abstraction

Attribute Name	Data Abstraction	Example values
PseudoID	Categorical	00000000 ~ 99999999
V1-15: Run 10 meters on flat surface, walk 5 stairs, ...	Ordinal	"Easily", "Slight difficulty", "Great difficulty", "I can not"
Score_ADL	Ordinal	0 ~ 45 Sum of the scores of V1 ~ V15 where "easily" is quantified as "0" and I can not is quantified as "4"
V17: How much pain have you experienced as a result of your disease in the past week?	Ordinal	0 ~ 100
V29: Do you have pain and/or swelling today in one of the joints listed below? LEFT	Categorical	"I have no pain in my left joints" OR any combination of "Fingers", "Wrist", "Elbow", "Shoulder", "Hip", "Knee", "Ankle", "Toes"
V38: Do you have pain and/or swelling today in one of the joints listed below? RIGHT	Categorical	"I have no pain in my right joints" OR any combination of "Fingers", "Wrist", "Elbow", "Shoulder", "Hip", "Knee", "Ankle", "Toes"
V40: Do you have pain and/or swelling today in one of the joints listed below?	Categorical	"I have no pain in my neck or back" OR any combination of "Neck" "Back"
V45: Have you PAST WEEK in the morning while waking up had STIFF JOINTS?	Boolean	"Yes", "No"
V46: How long did it take before the stiffness went away	Categorical	"", "15 min. or less", "15-30 min"
V52: Have you in past week experienced fever above 38°	Boolean	"Yes", "No"
V53: Rash	Boolean	"Yes", "No"
V55: Severity illness	Ordinal	1 - 100
V59: Illness description	Ordinal	"Remission", "Active", "Escalating"
V85: When you look at the last visit, how has your illness changed?	Ordinal	"Same", "worse"
V90: Do you use medicine for joint inflammation?	Boolean	"Yes", "No"
V96: Which medicine do u use at this moment?	Categorical	"NSAIDS", "Steroids", "Methotrexate", "
V100: Method of application methotrexate	Categorical	"Tablet", "injection",
V104: Experienced symptoms caused by medicine?	"Boolean"	"Yes", "no"
V105: What symptoms were bad?	Categorical	"Fever" AND/OR "Headache" AND/OR "Rash"...
V106: Other	text	
V111: Do you take your medicine on time's given by doctor?	"Boolean"	Yes / no

V112: How come?	Text	“”
V113: Other	Text	“”
V114: What medicine most difficult?	Text	“”
V115: Do you go to school?	Boolean	Yes / no
V116: Do you experience problems at school due to illness?	Boolean	Yes / no
V117: Which problems?	Text	“”
V118: Other	Text	“”
V119-128: Quality of life questions	Ordinal	“Never” .. “Always”
V129: ScoreQoL	Ordinal	0 = “best”, 30 = “worst”
V136: How do u feel right now?	Ordinal	0 = best, 100 = worst
V139: Would you be okay if your illness, as is, does not change in the coming months?	Boolean	Yes / no
V140: Can we use your data for scientific research?	Boolean	Yes /no

Table 1: Data Abstraction of JAMAR Patient dataset

Appendix B. Clinician Requirements Interview Protocol

The purpose of the interview is to gain a better understanding of the problems clinicians face during your daily clinics. The interview will be semi-structured; follow-up questions based on the answer given in the main questions. These follow-up questions are written down as sub questions, such as Q1.1. These function as support if the answer given to the main question is not sufficient. The clinician is asked for permission to record the interview for transcription so that their answers can be used optimally. Before pre-planned questions are asked, the clinician is asked several questions that were formulated based on the observations made during the user observation.

Q1: How do you typically use the results of the JAMAR in your clinical practice?

Q1.1: How do you use the results of the JAMAR outside of the 20-minute timeframe of an outpatient clinic?

Q1.2: Is it possible to gain an overview of the results over time? How do you do it with the current system?

Q1.3: Is it possible to gain an overview of multiple patients?

Q1.4: Is it possible to effectively give feedback to patients?

Q1.5: Besides what we have mentioned, is there anything else you are missing with the current overview of the JAMAR results?

Q2: Have there been systems used in the past to tackle potential limits of the current HiX system?

Q2.1: If so, to what extent were these systems successful and are there any reports about these systems?

Q2.2: Why are these systems no longer used today?

The clinician is presented with the JAMAR questionnaire that is to be filled in by patients. The goal is to gain an insight into the most valuable aspects of the JAMAR results according to the clinician.

Q3: Out of the 15 main questions of the JAMAR which answer would you consider the most

valuable in treating patients with JIA?

Q3.1: Could you arrange the five most important JAMAR questions?

Q3.2: Is there any aspect of the JAMAR questionnaire you do not value or intend to utilize while providing care?

For the most valuable JAMAR questions that were identified more specific questions are asked.

Q4: What would you like to be able to with these JAMAR results that you currently do not feel empowered to do?

Q4.1: Development of JAMAR results over time?

Q4.2: Comparing patients?

Q4.2.1: Comparing 2 specific patients or patients to the patient body?

Q4.3: Providing feedback on the JAMAR?

Q4.4: Adding information to JAMAR results through annotation?

Q4.5: Detecting patients through strongly differentiating values?

Q5: If a new system were to be build and implemented to better visualize the JAMAR, how would you like to be able to use this system?

Q5.1: Would primarily use it to discover new information?

Q5.1.1: Would you like to use it to generate new hypotheses or to verify new hypotheses?

Q5.2: Would you primarily use it to present data to third parties?

Q5.2.1: Which data and which third parties?

Q5.3: Would you primarily use it out of general curiosity about your patients or enjoyment?

Q6: How important is it for you to be able to enter new information to this new system?

Q6.1: Adding notes to generated visualizations

Q6.1: To record the current state of the system for later analysis?

Q7: To what extent would you like to be able to search for patients with this new system?

Q7.1: Would you primarily want to search for patients with the patient already being known or to browse through your patients without knowing the patient you are looking for?

Q8: How important would it be for you to identify patients with outlying values?

Q8.1: What sort of values would you like to be able to identify patients by?

Q9: How important would it be for you to be able to compare patients on measurements?

Q9.1: What sort of values would you like to compare patients by?

Q9.2: Would you prefer to compare specific patients, or a patient to the patient population?

Q10: Is there any information derived from the results of the JAMAR that is not appropriate to be shown to the patient?

Q10.1: Is it appropriate to view the development over time?

Q10.2: Is it appropriate to view a comparison to the patient body?

Q11: Are there any other elements that we have not discussed that you would like to see visualized in the new system?

Q11.1: Would you value predictions of future developments in the new system?

Q11.2: Are there elements besides the JAMAR results that you would like to see visualized?

Appendix C. Patient Requirements Interview Protocol

The goal of the interview is explained as learning about the experiences of patients with filling in the JAMAR questionnaire and to find out if there is a way that this process can be improved. The patient is informed about the fact that their name or any other data that could lead to an identification of the patient is not recorded, to provide maximum patient confidentiality. Before the interview starts, the patient is asked for written permission that the interview is recorded. The consent form details all the guarantees to the patient's confidentiality. Then, the patient is presented with a printed version of the JAMAR questionnaire and asked to narrate their experience with the JAMAR questionnaire.

Q1: How would you describe your experience with filling in the JAMAR questionnaire?

Q1.1: For how long have you been filling in the JAMAR?

Q2: What elements of the JAMAR questionnaire were confusing to you?

The patient is introduced to the idea of a doctor that has all patient data available in his memory and has analyzed all the patient data, he is aware of all your medical data throughout the past. Besides your patient data, the doctor knows of all the patient data of the patient population at the clinic and the doctor has a computer in front of him and all the possibilities of technology available to him.

Q3: Given these limitless possibilities, what would be interesting for you to discuss or learn from your doctor?

Q3.1: Would you value seeing your answers over time? If so, why?

Q3.2: See values that are very different from normal? If so, why?

Q3.3: Would you value seeing how you answer in comparison to the patient population? If so, why?

Appendix D Clinician Evaluation Protocol

The goal of the evaluation, testing the newly developed dashboard for JIA, is explained to the interviewed clinician. The clinician is asked for permission to record the evaluation. Upon receiving approval, the audio is recorded through two separate devices. Then, the clinician is asked to start up their electronic health record program, just like they would in their clinic. Instructions are provided to access the dashboard that is running during the evaluation on the second monitor. Once the desired setup is achieved, the structure of the evaluation is explained.

The evaluation starts with a tutorial of 10 minutes with the purpose of helping the clinician understand how the dashboard can be used. For this tutorial, the functionality of each user interface element is explained, starting with the patient number input field. The clinician is tasked to fill in a provided patient number to test this functionality. After the test patient is generated, the functionality of the dropdown menu with clinic dates is explained. Then, the clinician is tasked to select the most recent clinic, for which a JAMAR questionnaire is filled in. After the correct clinic is selected, the information provided in the narrative is explained as a summary of the JAMAR results. Then, the clinician is introduced to the variable selection dropdown menu. Each of the selectable variables and the way that they relate to the questionnaire is shortly described. Then, the clinician is tasked to select the functionality scale variable. The generated graphs and narrative are explained, and the functionality of the secondary graph that two variables have is explained. Then, the clinician is tasked to select all variables by double clicking on any variable in the legend. The graph will illustrate that there was one peaking yellow line at a specific date, and the clinician is tasked to visualize only that line by double clicking on the yellow line in the legend. Then, the clinician is asked to include two more lines in the line graph by selecting the variables in the graph legend. After the graphs are understood, the clinician is tasked to select the pain variable in the variable selection. The dashboard will visualize the data related to pain and the clinician is asked whether they can spot the most recent date where the patient experienced pain. Upon successful identification of the date, they are asked to select this date in the clinic selection menu, to learn in what joints the patient experienced pain at that date. Lastly the clinician is introduced to the raw data button to view the answers of the selected date in tabular format. The clinician is given the opportunity to ask any questions about functionalities of the dashboard that are unclear.

The evaluation itself takes place in two stages and the clinician is asked to narrate their thoughts in their own language through the *thinking aloud* protocol during the clinic. The first stage commences with a story about how the clinician is preparing for a clinic with a patient that the clinician will be seeing soon. The clinician is told that they are about to see a specific patient and that they are taking some time in the preparation of the clinic to study the results of the JAMAR questionnaire. This patient is a patient that the clinician will see in one of their next clinics in their real-world practice. The name of the patient is given to the clinician, and it is up to the clinician to copy the patient number from the electronic health record program to the patient number input of the dashboard. This will generate the relevant data for the patient so that the clinician can start their preparation. The clinician is told that to prepare for the clinic, they would like to learn how the patient is currently doing in their day-to-day life and what symptoms the patient is currently experiencing. Furthermore, the clinician would like to know if these aspects have been becoming better or worse over time.

During the second stage of the evaluation, the clinician is told to imagine that they are currently in a clinic with the patient that they prepared for. During this clinic, the clinician would like to allocate a few minutes to going over the results of the JAMAR with the patient. The clinician will go over the most recent JAMAR, pointing out parts of the patient reported outcomes that may be worth discussing with the patient. If the clinician does not generate any line-graphs by themselves, they are told that they are curious about the progression of the disease over time and to discuss this with the patient. Lastly, the clinician is interested in discussing to what extent the JAMAR questionnaire was filled in, and the effect this has on the generated dashboard.

After the experiment, the following questions are asked to the clinician. Observations that were made during the experiment, are pointed out to the clinician, following the *critical incident* technique used in the requirements interview. The questions are asked in Dutch.

Q1: What did you think about using the dashboard to learn about the patient reported outcomes?

Q2: A specific incident is pointed out where the dashboard seemed to provide a valuable addition to understanding the patient reported outcomes, and the clinician is asked to what extent he felt supported by the tool at that moment.

Q2.1: Are the line graphs useful to learn about the progression of the patient?

Q2.2: Is the narrative summary effective at providing a summary of the last JAMAR submission?

Q2.3: Is the narrative effective at providing more detailed information based on variable and clinic date selection?

Q2.4: Is the clinic dates overview with corresponding JAMAR dates useful to discuss the extent to which patients fill in the questionnaires?

Q3: A specific incident is pointed out where clinician experienced an issue with the dashboard, and the clinician is asked whether he felt sufficiently supported by the tool at that moment.

Q3.1: How could the tool be changed to support the physician more effectively?

Q4: Would you use the revised version of this dashboard in your next clinic?

Q4.1: During the preparation of the clinic, or during the clinic itself?