

# A Visual Analytics Dashboard for Mental Health Therapists

*Supporting Guided Internet-based Cognitive  
Behavioral Therapy*

Nikolai Alexander Grieg

Master Thesis in Software Engineering at  
Department of Computing, Mathematics and Physics,  
Western Norway University of Applied Sciences

Department of Informatics,  
University of Bergen

Supervisors: Svein-Ivar Lillehaug and Yngve Lamo

June 2019



Western Norway  
University of  
Applied Sciences





## Abstract

Mental health disorders are a significant problem throughout the western world. Cognitive Behavioral Therapy (CBT) is a common treatment form for these disorders. Internet-based Cognitive Behavioral Therapy (iCBT) provides higher efficiency than face-to-face CBT, without reducing the efficacy of the treatment. While iCBT has been shown to work well on a small scale, the amount of data each therapist need to consider, grows rapidly with the number of patients. This thesis examines the data used for iCBT, and presents possible solutions for gaining overview of this data through visual analytics and dashboard techniques.

We designed and developed an artifact through the design science methodology, in order to implement these techniques. The main evaluation of the artifact consisted of semi-structured interviews with experts in psychology and psychiatry. We present methods for automatic processing of standardized mental health screening questionnaires, as well as a novel way of visualizing these questionnaires over time. We also discuss the scalability of a dashboard application using the HL7 FHIR standard for representing screening questionnaires and numerical observational data. Finally, we present identified problems and further work on structuring data for analytics in iCBT.

## Sammendrag

(this is a Norwegian translation of the abstract)

Psykiske lidelser er et betydelig problem i den vestlige verden. Kognitiv Atferdsterapi (CBT) er en vanlig behandlingsform for denne typen lidelser. Internett-basert Kognitiv Atferdsterapi (iCBT) gir økt effektivitet sammenlignet med ansikt-til-ansikt CBT, uten å redusere effekten av behandlingen. Selv om iCBT virker bra på liten skala, vokser mengden data hver terapeut trenger å forholde seg til, raskt med tanke på antall pasienter. Denne oppgaven undersøker data som blir brukt i iCBT, og presenterer mulige løsninger for å få oversikt over disse gjennom visuell analyse og dashboard-teknikker.

Vi designet og utviklet en artefakt ved hjelp av design science metodikken, for å implementere disse teknikkene. Hovedevalueringen av artefakten besto av semi-strukturerte intervjuer med eksperter i psykologi og psykiatri. Vi presenterer metoder for automatisk prosessering av standardiserte spørreskjema, i tillegg til en ny måte å visualisere disse spørreskjemaene over tid. Vi drøfter også skalerbarheten av en dashboard applikasjon som bruker HL7 FHIR standarden for å representere spørreskjemaer og numeriske observasjonsdata. Til slutt presenterer vi de identifiserte problemene og videre arbeid for strukturering av data for analyse i iCBT.

## Acknowledgements

I would like to thank my supervisors Svein-Ivar Lillehaug and Yngve Lamo for their advice and guidance throughout the process of working with this thesis. I would also like to thank the experts in psychology and psychiatry who helped me evaluate my work. Finally I would like to thank my family for all the support they gave me these two years.



# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Research Questions . . . . .	2
1.3	Research Methods . . . . .	2
1.3.1	Design Science . . . . .	3
1.3.2	Problem identification . . . . .	3
1.3.3	Objectives for a solution . . . . .	3
1.3.4	Ethical declaration . . . . .	5
1.4	Thesis Structure . . . . .	5
<b>2</b>	<b>Background</b>	<b>6</b>
2.1	Cognitive Behavioral Therapy (CBT) . . . . .	6
2.2	Internet-based Cognitive Behavioral Therapy (iCBT) . . . . .	6
2.2.1	Data in iCBT . . . . .	8
2.3	Dashboards and Visual Analytics . . . . .	11
2.4	Fast Healthcare Interoperability Resources (FHIR) . . . . .	13
2.5	Related Work . . . . .	14
2.5.1	Practical Visualization of Multivariate Time Series Data in a Neonatal ICU . . . . .	15
2.5.2	Multidisciplinary Intervention in Patients with Muscu- loskeletal Pain: a Randomized Clinical Trial . . . . .	16
2.5.3	Supporting Clinical Cognition: A Human-Centered Ap- proach to a Novel ICU Information Visualization Dashboard . . . . .	17
2.5.4	A dashboard-based system for supporting diabetes care . . . . .	17
2.6	Section summary . . . . .	18
<b>3</b>	<b>Method and Design</b>	<b>20</b>
3.1	Design process . . . . .	20
3.2	Artifact description . . . . .	21
3.2.1	Master View . . . . .	21
3.2.2	Detail View . . . . .	24
3.3	Psychometric screening data as time series . . . . .	27
3.4	Visualizing multidimensional time series . . . . .	30
<b>4</b>	<b>Implementation</b>	<b>32</b>
4.1	Tools and frameworks used during development . . . . .	32
4.1.1	Matplotlib and Seaborn . . . . .	32
4.1.2	Dash . . . . .	32
4.1.3	SMART on FHIR JavaScript client . . . . .	32
4.1.4	SMART on FHIR Python client . . . . .	33
4.2	Tools and frameworks in the final artifact . . . . .	33
4.2.1	ASP MVC . . . . .	33
4.2.2	MongoDB . . . . .	33
4.2.3	HAPI JPA . . . . .	33
4.2.4	Synthea . . . . .	34
4.2.5	Highcharts and D3 . . . . .	34
4.3	System Architecture . . . . .	34
4.3.1	Update flow . . . . .	35

4.4	Mental health data generation . . . . .	37
4.5	Status variables . . . . .	38
4.5.1	Replaceable functions . . . . .	39
4.6	Scalability . . . . .	41
4.6.1	Scaling measures to improve user-experienced latency . .	42
4.6.2	Possible scaling measures to improve update time . . . . .	42
4.7	Security considerations . . . . .	44
<b>5</b>	<b>Demonstration and Evaluation</b>	<b>45</b>
5.1	User evaluation through tests and semi-structured interviews . .	45
5.1.1	Interview results . . . . .	46
5.2	Scalability measurements for evaluating system performance . . .	47
5.3	Heuristic evaluation with dashboard design guidelines . . . . .	50
<b>6</b>	<b>Findings and Discussion</b>	<b>53</b>
6.1	Contributions to the knowledge base and answers to the research questions . . . . .	53
6.2	Contributions to the domain . . . . .	55
6.3	Validity . . . . .	56
6.3.1	Internal validity . . . . .	56
6.3.2	External validity . . . . .	57
6.4	Reflection over tools and frameworks . . . . .	57
6.5	Communication . . . . .	59
<b>7</b>	<b>Conclusion and Further Work</b>	<b>60</b>
7.1	Conclusion . . . . .	60
7.2	Further Work . . . . .	60



## List of Figures

2.1	Excerpt of the MADRS questionnaire [23] . . . . .	7
2.2	From the left: Line chart, bar chart, pie chart . . . . .	12
2.3	Metaphor graphic for displaying physiological measurements for a single patient from [46] . . . . .	13
2.4	Animated spider chart paused, from Lehmann et. al. [57] . . . . .	16
2.5	Spider chart in ISIVET . . . . .	16
2.6	Temporal displays for biometric variables for a single patient in an ICU with MIVA 1.0 . . . . .	17
2.7	Selected visualizations from dashboard for a patient population in Dagliati et. al. [60] . . . . .	18
3.1	Master View . . . . .	22
3.2	Pie- and Bar chart in Master View (1 and 2) . . . . .	23
3.3	Warnings (3), and tools for search and filtering of table (4) . . . . .	23
3.4	Detail View . . . . .	24
3.5	Picture and background information (1) and selectors (2) . . . . .	25
3.6	Line charts (3) . . . . .	26
3.7	Example of a HL7 FHIR QuestionnaireResponse resource for the MADRS questionnaire . . . . .	29
3.8	Interactive Spider Chart . . . . .	31
4.1	Architecture model . . . . .	34
4.2	Sequence diagram for the update flow . . . . .	36
4.3	Code snippet for simulating realistic answers to questionnaires . . . . .	37
4.4	Example of a HL7 FHIR Observation resource measuring a patients sleep duration for one night . . . . .	38
4.5	The interfaces provided for calculating status variables . . . . .	40
5.1	Interview setup . . . . .	46
5.2	Code snippet for requesting all QuestionnaireResponse resources for one patient . . . . .	48

## List of Tables

2.1	All questionnaires given to patients in iCBT treatments at eMeistring (from eMeistring internal documentation, October 2017) . . . . .	9
3.1	Answers to the MADRS questionnaire in tabular format . . . . .	28
3.2	Multidimensional time series example . . . . .	28
5.1	Total resources retrieved in one update . . . . .	48
5.2	Single patient measurements . . . . .	49
5.3	Measurements across all patients . . . . .	49
5.4	Single patient measurements, with paging . . . . .	49
5.5	Measurements across all patients, with paging . . . . .	50

## Abbreviations

- API = Application Programming Interface
- BAI = Beck Anxiety Inventory
- BSON = Binary JSON
- CBT = Cognitive Behavioural Therapy
- CDS = Clinical Decision Support
- CRUD = Create Read Update Delete
- D3 = Data Driven Documents
- EHR = Electronic Health Record
- FHIR = Fast Healthcare Interoperability Resources
- GUI = Graphical User Interface
- HAPI = HL7 API
- HL7 = Health Level Seven
- iCBT = Internet-based Cognitive Behavioural Therapy
- ICU = Intensive Care Unit
- IDE = Integrated Development Environment
- IoT = Internet of Things
- JPA = Java Persistence API
- JSON = JavaScript Object Notation
- MADRS = Montgomery And Aasberg Depression Rating Scale
- MADRS-S = Montgomery And sberg Depression Rating Scale (Self-filled)
- MVC = Model View Controller
- PHQ-9 = Patient Health Questionnaire-9
- QR = QuestionnaireResponse
- REST = Representational State Transfer
- SMART = Substitutable Medical Applications, Reusable Technologies
- XML = eXtensible Markup Language



# 1 Introduction

## 1.1 Motivation

The problem of mental health disorders have gained focus over recent decades [1]. Mental health disorders affect more than a third of the European population, as of estimates from 2010 [2], and it is expected that more than 50% of the population in middle and high income countries will suffer from at least one mental disorder at some time throughout their lives [3]. A Swedish study from 2019 shows increasing rates of mental health disorders in youth, with 41% of girls and 30% of boys having psychosomatic problems, and argues for a need to start suitable treatment as fast as possible after diagnosis [4]. Furthermore, mental health disorders has been calculated to account for 13% of global disability-adjusted life-years (DALYs) [5].

Amongst the consequences of mental health disorders, in addition to the effects on the patient's quality of life, are major economic costs for the society. Accounting for both direct healthcare costs, and indirect costs such as productivity loss, mental health disorders are estimated to account for higher economic costs for the society than chronic somatic diseases, such as cancer or diabetes [3].

The most frequent mental health disorders in Europe in 2010 were anxiety, insomnia, depression and somatoform disorder [2]. Cognitive Behavioral Therapy (CBT) is a well documented treatment form used to treat patients with any of these disorders [6][7][8].

While CBT has shown positive effects on patient outcome [9], each therapist does only have the capacity to treat a handful of patient each day with this method. There is a "treatment gap" for mental health disorders, partially due to lack of accessible treatment options [10]. Internet-based Cognitive Behavioral Therapy (iCBT) has emerged in an effort to make CBT more accessible and efficient. Therapist guided iCBT has shown the same efficacy as regular CBT for some patient groups [11], while reducing the time required per patient for the therapist. An evaluation of the treatment given at eMeistring in Bergen, Norway suggest therapists working with iCBT can help 10-12 patients per day, while therapists providing face-to-face CBT had 3-4 consultations per day [12].

While iCBT has provided promising results so far, there is still room for improvements. The treatment relies on an Internet-based platform, both for patient-therapist communication and for delivery of the treatment programmes. The use of this platform provides potential for data analytics to support the therapists in their work, an approach which to our knowledge has not been applied for iCBT treatments yet.

Better presentations of the data can improve therapist workflow by removing the need to manually consider each data point. Providing automatic processing of complex data sets can further optimize therapists performance. By gaining a better representation of the patient state through the available data, therapist

can tailor the treatment better towards the single patient. Therapists can also be better equipped to prioritize within their group of patients. Combined, this can lead to both higher efficiency and efficacy of therapist guided iCBT.

This situation presents an opportunity to improve the treatment by applying data visualization and Clinical Decision Support (CDS) tools. These techniques can provide insight both within the context of a single patient, and a their group of patients. Therapists can identify trends in the data for a single patient, which can provide opportunities to customize the treatment to the patient. For a group of patients, a fast overview based on the data can provide faster help, and more of the therapist's time, to the patients needing it the most. This resembles the medical technique of triage, where treatment providers assign each patient degrees of urgency depending on their need for care [13], although on a different time scale. A better overview of a group of patients with summary or status variables presented to express the need of each individual patient might also contribute to help the therapist organize her work day in a better way.

## 1.2 Research Questions

The overall research question of this thesis is how to design and implement an artifact to deliver visual analytics and decision support for therapists in iCBT. Specifically, we will explore the following research questions:

- RQ 1** What data representation of mental health data in iCBT can facilitate analytics?
- RQ 2** Using the data representation from RQ1, how can the data be visualized for therapists to gain actionable insights into the state of a single patient or a population of patients?
- RQ 3** How can decision support functionality be implemented to assist therapists in understanding each patients required urgency of treatment in a population of patients, based on the data from RQ1?
- RQ 4** How will a dashboard implementing the techniques from RQ2 and RQ3 scale when using the HL7 FHIR format for medical interoperability?

## 1.3 Research Methods

As this thesis aims to solve particular problems of practical iCBT, we chose the design science methodology. Other research methods could also be applied for this problem, such as the action research methodology or a mixed methods approach. We found the process of design science, as outlined below, most applicable to our situation. The process of designing an artifact to solve a real problem is the core of the research methodology. With the goal of scaling the iCBT treatment with respect to mental health data handling, a prototype application has been designed.

### 1.3.1 Design Science

The Design Science Research Methodology for Information Systems Research, as described by Pepper et. al. [14], prescribes six steps to take when applying this methodology. These steps include creating an artifact to solve a relevant problem in practice. The six activities are defined as:

1. Problem identification and motivation
2. Define the objectives for a solution
3. Design and development
4. Demonstration
5. Evaluation
6. Communication

Motivation is described in the first section, following the problem relevance guideline defined by Hevner et. al. [15]. This guideline states that "the objective of design-science research is to develop technology-based solutions to important and relevant business problems". Section 1.1 discussed the importance and relevance of the problems in iCBT. The following two sections will describe the problem identification process, and the objectives for the design science solution. The design and development of the artifact is presented in chapters 3 and 4. The demonstration and evaluation activities are presented in chapter 5. To a large degree this thesis represents the communication activity, which is elaborated further in section 6.5. Also, we plan to summarize the results of this work in a paper intended for scientific publication.

### 1.3.2 Problem identification

To explore the idea of using data analytics in iCBT, we initially had a meeting at the eMeistring [16] clinic in Bergen. In this meeting, technical challenges were identified as possible points of improvement in the delivery of the current iCBT treatment program, eMeistring (eCoping in english). Some of these challenges are also described by Folker et. al. [12] in their analysis of iCBT treatments in Europe. The problems with the platform are identified as major issues for therapists and patients. The platform does not allow for interactive content, and making changes to the platform is slow and difficult.

The initial meeting at the eMeistring clinic consisted of an unstructured conversation with the therapists involved in iCBT. A lack of tools for visualizing and presenting patient data in their current systems was a highlighted point from this meeting. Our initial plan was to work with these therapists throughout the design and development process. Unfortunately this was not possible due to a lack of human resources at the eMeistring clinic.

### 1.3.3 Objectives for a solution

To solve the problems identified in section 1.3.2 we formed objectives for an artifact. During this process, we performed a literature survey to identify possible similar systems. Through searching the scientific literature databases Google

Scholar and Medline, no clinical dashboard applications for use in iCBT, nor Internet-based therapy in general, were discovered. To identify potential methods for solving the problems identified at the eMeistring clinic, we performed literature surveys on the use of dashboards and visualizations in other medical domains. We also considered the use of visual analytics and dashboards in general. Some relevant projects are presented in section 2.5.

As the identified problems in iCBT can be condensed to a need for better utilizing the electronically available data, some objectives for implementing the artifact can be created based on this. Through better utilization of the data, the therapists can make more informed decisions. This can in turn provide higher quality of care or increase the availability of care, through potentially more efficient therapists. As the actual patient data used in health IT applications are sensitive, it can be both complicated and unpractical to test a prototype on real patient data. However, given descriptions of the data used, some objectives can be described and tested based on generated data. As the data is generated, fitting the data to a standard for interoperability ensures the application can accept real data as input. The process of developing a solution should:

1. be based on standardized formats for representing patient data to enable further development of the application, and testing on real patient data.
2. identify ways to present the data to enhance the users workflow and give actionable insights.
3. provide solutions for identifying how patients undergoing iCBT are progressing.
4. scale the iCBT treatment to enable more patients per therapist, with respect to analysis of mental health data.

To solve these objectives, we chose an application with dashboard functionalities for the implementation of the artifact. As a full implementation of a clinical dashboard is a massive undertaking, the scope of the implementation was constrained to some of the functionalities for such a system. The scope of the artifact development was constrained in the following ways:

- The application supports a subset of the potentially useful data types in iCBT. Currently, therapists in iCBT make use of significant amounts of unstructured data, as described in section 2.2.1. As one of the goals for the artifact is to conform to a standard for medical data interoperability, the application supports data conforming to certain HL7 FHIR resources. Specifically, we focused on questionnaires and quantitative observations as the data types for the artifact. These data types were determined to provide the most value within the scope of the project. The FHIR resources are described further in section 2.4.
- The artifact does not address any security issues. For an actual clinical application, securing the patient data would be of uttermost importance. As there were multiple other aspects of the artifact which are more important for answering the research questions, these were prioritized. The artifact was developed with security in mind for future development, and considerations with respect to this is described in section 4.7.



- The artifact was primarily designed towards Internet-based treatment of depression. As such, the data types used for the therapist guided iCBT treatment of depression at eMeistring set the baseline for the development. The use of standardized screening questionnaires are used heavily here, as seen in table 2.1.

#### 1.3.4 Ethical declaration

As no real patient data was used for building the artifact in this thesis, an ethics approval was not necessary. For the semi-structured interviews, both interviewees verbally consented to being interviewed on video.

### 1.4 Thesis Structure

An outline of the thesis is presented here. The most important topics of each chapter are presented briefly.

- **Chapter one** presented the problem identification and motivation for the thesis, as well as the objectives for a solution.
- **Chapter two** contains theoretical background information and related work. Cognitive Behavioural Therapy (CBT) and Internet-based CBT (iCBT) are described. Further, dashboards and visual analytics are presented. Next, the HL7 FHIR standard is described. Finally, related work and systems with similar capabilities are discussed.
- **Chapter three** presents the artifact design process and the final artifact design. The design process and considerations relating to the design of the GUI is presented. The final GUI is also explained.
- **Chapter four** presents a detailed description of the implementation of the artifact. Tools used to develop the artifact is presented, and the system architecture is described. The generation of FHIR resources follows. The implementation of status variables and decision support is then presented. Lastly, scalability and security of the artifact is discussed.
- **Chapter five** presents the demonstration and evaluation activities of the design science research. The chapter first presents semi-structured interviews with experts in psychology and psychiatry to evaluate the artifact. Following this, empirical measurement of run time is presented. Lastly, we evaluate the artifact towards dashboard design guidelines from the literature.
- **Chapter six** presents the research findings, and provides a discussion about the design science research process. Contributions to the knowledge base and the problem domain is presented here. The validity of the research is then discussed, followed by a reflection over the tools used in the artifact, and the communication activity of design science.
- **Chapter seven** concludes the thesis and presents further work.



## 2 Background

This chapter will give an introduction towards patient data, and the use of such data in CBT and iCBT, along with techniques and practices relevant for solving the design science objectives. The first two sections provide a description of CBT and iCBT. A brief introduction to dashboards and visualization theory follows. The HL7 FHIR standard for medical interoperability is then presented, as the data types in this standard are central for the artifact in the thesis. Finally, related work is presented, consisting of systems from other medical domains used to analyze and visualize patient data.

### 2.1 Cognitive Behavioral Therapy (CBT)

Cognitive Behavioural Therapy is a form of psychotherapy applicable to a wide range of mental health problems, including depression, anxiety and substance abuse [17]. In CBT, a therapist works with her patients suffering from mental diseases, to improve their thinking- and behavioural patterns [18]. In many modern CBT programmes the patients will fill questionnaires in order to track their progress over the course of the treatment. The questionnaires used in practice are clinically validated and demonstrates psychometric properties [19]. In other words, the questionnaires can be used to accurately determine the patients state, and progress over time. Some examples of questionnaires used for CBT are MADRS [20], PHQ-9 [21] and BAI [22]. These questionnaires are used in different sub-disciplines within mental health therapy. MADRS and PHQ-9 are examples of questionnaires to screen for depression, whereas BAI is used to screen for anxiety. Figure 2.1 shows questions three and four of the MADRS questionnaire.

### 2.2 Internet-based Cognitive Behavioral Therapy (iCBT)

Internet-based CBT (iCBT) can take the form of guided or unguided self-help programs [24]. In guided self-help programs, a therapist will guide the patient through the treatment. This guidance consists of electronic asynchronous communication such as email or messages through a patient portal. Some treatment programs offer mixed CBT and iCBT, referred to as blended care [25]. These programs consists of iCBT supported by face-to-face consultations, as opposed to only electronic communication. Unguided self-help programs are fully automated and involves no contact between patient and therapist.

There is some disagreement about whether guided self-help programs provide higher efficacy. Studies on depression [24], perfectionism [26], and binge eating disorder [27] has indicated that guided self-help programs can provide a better reduction in symptoms compared to unguided therapy programmes. For studies in pain management [28] and social phobia [29], no differences were found between patient groups with guided- and unguided self-help. Guided self-help provides additional opportunities with regards to adapting the treatment to the patient. For this form of iCBT, there is a potential for using technology and data analytics to improve the treatment. Because of this, guided self-help will be the iCBT format addressed in this thesis.

<b>3. Inner tension</b> Representing feelings of ill-defined discomfort, edginess, inner turmoil, mental tension mounting to either panic, dread or anguish. Rate according to intensity, frequency, duration and the extent of reassurance called for.	
<b>0</b> = Placid. Only fleeting inner tension.	<input type="checkbox"/>
<b>2</b> = Occasional feelings of edginess and ill-defined discomfort.	<input type="checkbox"/>
<b>4</b> = Continuous feelings of inner tension or intermittent panic which the patient can only master with some difficulty.	<input type="checkbox"/>
<b>6</b> = Unrelenting dread or anguish. Overwhelming panic.	<input type="checkbox"/>

<b>4. Reduced sleep</b> Representing the experience of reduced duration or depth of sleep compared to the subject's own normal pattern when well.	
<b>0</b> = Sleeps as normal.	<input type="checkbox"/>
<b>2</b> = Slight difficulty dropping off to sleep or slightly reduced, light or fitful sleep.	<input type="checkbox"/>
<b>4</b> = Moderate stiffness and resistance	<input type="checkbox"/>
<b>6</b> = Sleep reduced or broken by at least 2 hours.	<input type="checkbox"/>

**Figure 2.1:** Excerpt of the MADRS questionnaire [23]

A main focus in iCBT is for patient to perform exercises and activities organized through modules, where each module has its focus on some particular topic. These modules are represented one by one as homework assignments for the patients [30]. We can consider the modules offered in the depression treatment at eMeistring as examples of this. The treatment is intended to span over 14 weeks through eight treatment modules [31]. The various modules addresses the following topics:

- Module 1: Introduction and information about CBT.
- Module 2: Planning activities. Here, patients are taught to think of activities as "plus-activities" and "minus-activities".
- Module 3: Achieving balance in the day. This module focuses on increasing the amount of plus-activities.
- Module 4: Reducing the amount of minus-activities.
- Module 5: Discovering negative thoughts.
- Module 6: Finding strategies for dealing with the negative thoughts.

Module 7: Sleep, and handling sleep problems.

Module 8: Planning the future. This module recaps the previous modules.

Each of the first six modules last one week. Module seven spans five weeks, and module eight concludes the treatment program. As seen in table 2.1, the patients fill in the MADRS questionnaire every week. Additional questionnaires are also filled, depending on which treatment program the patient is in. Table 2.1 contains the full description on which questionnaires the patients in eMeistring are given. All the questionnaires except for "Background" and "Evaluation" are mental health screening questionnaires. Questionnaires in Norwegian are used for these treatment programmes.

Amongst the differences between iCBT and CBT is the asynchronous communication introduced in iCBT. In CBT, the patient and therapist meet physically to talk about the patients problems, while in iCBT the therapist will provide feedback on the patients exercises when available to do so. Typically in eMeistring, a patient will receive feedback from his therapist once or twice a week [32].

The nature of the asynchronous patient-therapist communication opens up for new therapist workflow options. Data analytics and visualizations has the potential to enable therapists to choose their workflow based on available information, as opposed to having a strictly synchronous workflow. Determining workflow based on prioritization, rather than a predetermined sequence, has successfully been applied in other medical domains [33]. Through aggregation and prioritization of the information presented for clinical consideration, eligible individuals can be identified within a population to better meet their needs [34].

The standardized questionnaires used in CBT are central for monitoring patient progress in iCBT. For the questionnaires MADRS and BDI, the psychometric properties of the questionnaires are not affected by medium (internet/-paper) [35]. This implies that the questionnaires are equally effective for CBT and iCBT. Therapists will not have direct contact with their patients when providing guidance. This increases the importance of monitoring patient progress through the data reported in these questionnaires.

### **2.2.1 Data in iCBT**

In eMeistring, therapists will usually start their work session by looking through data for each of their patients (see question 10 of the interview in Appendix A). Therapists will consider one patient at a time. The current platform also provides limited overview within the single patient. Questionnaires describing the current state of the patients are available to the therapists through a pdf file [31], or as a bar chart over the single questionnaire. For therapists to get an overview of a patient's progress, they will need to compare multiple files/views for any patient. Therapists will also need to rely on memory in order to gain an overview, resulting in significant cognitive load.

Time in programme	Day	Social Anxiety	Panic Disorder	Depression
PRE-module	PRE	MADRS-S	MADRS-S	MADRS-S
		SPS	BSQ	GAD
		SIAS	ASQ	PHQ
		SE-Scale	MI	IS
		TIC-P	SE-Scale	SE-Scale
		Background	TIC-P	Background
			Background	
Module 1	Started treatment (0)	MADRS-S	MADRS-S	MADRS-S
		C-Scale	C-Scale	C-Scale
Module 2	Day 7	MADRS-S	MADRS-S	MADRS-S
		SPS	BSQ	
	Day 14	MADRS-S	MADRS-S	MADRS-S
		SPS	BSQ	
Module 3	Day 21	MADRS-S	MADRS-S	MADRS-S
		SPS	BSQ	
		SIAS	ACQ	
			MI	
Module 4	Day 28	MADRS-S	MADRS-S	MADRS-S
		SPS	BSQ	
Module 5	Day 35	MADRS-S	MADRS-S	MADRS-S
		SPS	BSQ	
Module 6	Day 42	MADRS-S	MADRS-S	MADRS-S
		SPS	BSQ	
		SIAS	ACQ	
			MI	
Module 7	Day 49/56/63/70/77/84	MADRS-S	MADRS-S	MADRS-S
		SPS	BSQ	
Module 8	POST			MADRS-S
				GAD
				PHQ
				IS
				TIC-P
				Evaluation
Module 9		MADRS-S	MADRS-S	
		SPS	BSQ	
		SIAS	ACQ	
		TIC-P	MI	
		Evaluation	TIC-P	
			Evaluation	

**Table 2.1:** All questionnaires given to patients in iCBT treatments at eMeistring (from eMeistring internal documentation, October 2017)

For patients completing iCBT treatment programs, each patient will produce over 20 filled questionnaires (see table 2.1). With multiple patients per therapist, the amount of data will quickly become hard to process in pdf format. Therapists working with iCBT one day a week at eMeistring will have 15-20 patients, and therapists working full time in iCBT will have 50-60 patients. Without further processing of the data, most of this data will only be used for a short period after it has been gathered. Getting an overview across all patients will also be difficult when the questionnaires have to be processed manually.

While important, the standardized screening questionnaires are not the only data types considered in iCBT. Patients in eMeistring will have an initial face-to-face consultation before they commence the iCBT treatment program. This initial consultation is used, among other things, for diagnosing the patient and assigning an appropriate treatment program. During the consultation, therapists will gain multiple qualitative observations. Examples of such observations are patients talking slow, moving slow, or having cognitive impairments. These data types are not easy to process and visualize, and often also hard to record in a structured format.

One of the main sources of information generated by the patients is the exercises in the treatment modules. The answers/results of these exercises are integral to the treatment process, and are communicated to the therapists. As the exercises are not standardized and often based on free text, this data is currently difficult to process. Even though the exercises are important for iCBT, because of the unstructured nature they are not considered for the artifact in this thesis. Structuring and utilizing this data better could be an important and interesting topic for further research.

Other quantitative measurements than the standardized screening questionnaires can also be used in iCBT. Examples of potentially useful data sources are IoT data streams and application monitoring. Most of the data from these sources can be visualized as single dimension time series, for example by line charts.

An IoT data stream consists of high velocity data from a physical sensor. Smart watches or smart phones used for monitoring vital signs are common sources for these data streams. The raw data from the sensors can often be converted to new information. A continuous measurement of pulse could for example be converted into estimates for how many hours a person slept during the night. Converting the data this way also make the data easier to store in health records, as the aggregated data will have significantly lower velocity.

Data from sensors can be used in addition to the self-reported values by the patients. The values reported by patients might be inaccurate for a number of reasons. Possible error sources includes the patient forgetting to report within the expected timespan, making inaccurate estimates, or purposely not answering correctly. Data from sensors are subject to measurement error, but will generally produce reliable results. Sensor data can also give answers the patients would

not be able to answer. Sensors can for example be used to track the duration of REM (Rapid Eye Movement) sleep a patient received during a night.

Application monitoring will provide data on how a computer system is used in practice. Logging the actions of patients can provide information such as when a patient last logged in, frequency of logins, or amount of time spent on certain activities. In the context of iCBT, this information can be used to track adherence to the treatment program, as program adherence can be linked to patient outcome [24]. Tracking adherence to the program can provide actionable information, both on a patient- and population level for the treatment program.

To summarize: In the current iCBT treatment at eMeistring, therapists have to manually consider all the data for a patient. Relevant data types includes standardized screening questionnaires, qualitative observations, patient exercises and application monitoring. IoT data streams can also be used if the necessary hardware is available. The qualitative observations are not recorded in a structured manner, and the patient exercises are free-text based. As these data types are hard to process and visualize, the questionnaires and other quantitative temporal data are considered in this thesis.

### 2.3 Dashboards and Visual Analytics

Thomas and Cook defines visual analytics as the science of analytical reasoning supported by interactive visual interfaces [36]. Keim et. al. [37] further defines visual analytics to combine "automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets". While visual analytics includes visualization, it also considers the disciplines of decision-making, human factors and data analysis.

By leveraging multiple disciplines, visual analytics aims to reduce the effect of information overload, and achieve higher utilization of large data sets. Procedures to automatically process data can be implemented, and when further automatic analysis is intractable the result can be integrated with visualization and interaction techniques [37]. The techniques of visual analytics be combined with dashboard techniques to create tools for continuous data processing and visualization.

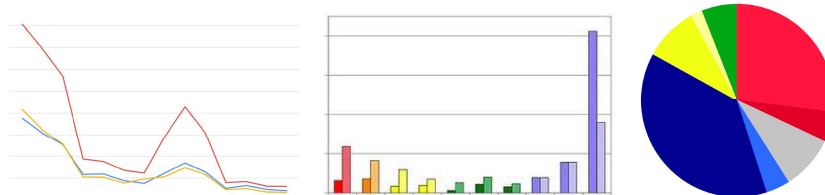
A dashboard application can be defined as a decision support tool capable of using multiple databases and providing visual representations of important performance indicators in a single report, often visually similar in format to a car's dashboard [38]. Dashboard applications can be further categorized into sub categories. Among these are: Clinical dashboards, performance dashboards [39], and administrative dashboards [40]. For this thesis, the main focus will be on clinical dashboards. Techniques for designing the user interface and visualizing data, however, are applicable across the dashboard sub categories.

The use of dashboards originated in business domains [41]. Examples of these domains are airlines [42], digital manufacturing [42], and automotive industries



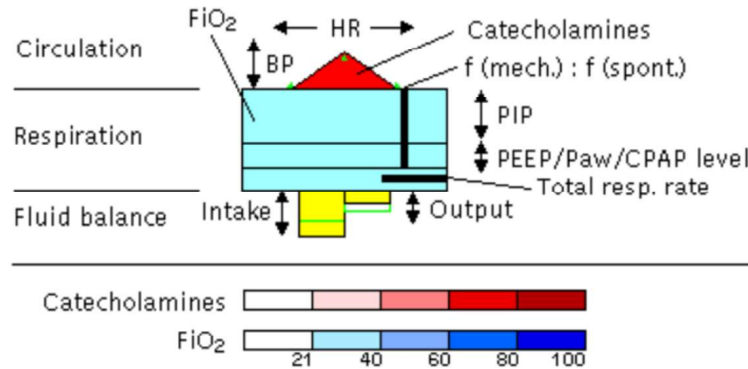
[43]. Dashboard techniques for data processing and visualization have also been proven valuable within health domains. Common places to see dashboards in healthcare are within the Intensive Care Unit (ICU) and for administrative tasks. Clinical dashboards show increasing efficiency, quality, safety and clinician satisfaction in some situations [44]. The visualizations used in ICUs focus on providing fast and accurate information to medical staff, with the goal of reducing cognitive load and provide better care. Administrative dashboards are often used on a population of patients, for example to visualize patient flow in a hospital [45]. In both of these cases, dashboard applications process and presents data that might have too high velocity or volume for a single person to interpret.

A wide range of visualization techniques are used to create clinical dashboards. Simple charts are strong default choices, as most users are used to these visualizations from previous experience. The users will require less time to get used to the visualization method, and can rapidly understand the information in the data. Examples of these charts are line charts, pie carts, and bar charts, as displayed in figure 2.2.



**Figure 2.2:** From the left: Line chart, bar chart, pie chart

More complex visualizations can be used when basic visualizations are not sufficient for the data at hand. These more complex visualizations can require some training to fully utilize, but they have the potential to provide faster information retrieval or contribute to the discovery of less apparent patterns. Examples of such visualizations are metaphor graphics (composite icons) [46], animations, or 3D objects. Metaphor graphics can be used for ICU dashboards [46][47], as they can provide an overview of a patients status at a glance. Figure 2.3 shows a metaphor graphic designed to visualize data in an ICU.



**Figure 2.3:** Metaphor graphic for displaying physiological measurements for a single patient from [46]

Making visualizations interactive aids the search for information. A commonly used visualization principle, known as Shneiderman’s mantra, is ”overview first, zoom and filter, then details-on-demand” [48]. Adding options for drill-down, filtering, etc. can make a visualization more powerful, as a user has more control to adapt the visualization to her current information needs.

## 2.4 Fast Healthcare Interoperability Resources (FHIR)

The HL7 FHIR standard [49] provides a common data format for health data. This is achieved by dividing data into logical ”resources” where each resource represents a concept. The resources can be expressed in formats such as JSON or XML, and communicated over REST. The intention of the standard is to provide interoperability between health IT systems. For dashboards gathering data from various external sources, this is important. As the primary function of dashboard applications is to present data, little to none of the data is created by the application itself. Consequently, adapting a common standard for the data used in the dashboard can simplify both development and maintenance of this type of system. Using a standard to ensure interoperability is also a common requirement in modern health IT applications. The US Food and Drug Administration (FDA) supports and promotes the use of standards for medical interoperability [50]. The Norwegian Directorate of eHealth recommends using the HL7 FHIR standard for implementing APIs in health IT applications [51].

A FHIR resource describes an object of data, generally on the resource server. All standard resources are described at the HL7 FHIR web page. A resource will have some required fields, and often some optional fields. Amongst the resources relevant for mental health data are the Patient, Questionnaire, Questionnaire-Response resources. The generic Observation resource also provides an option to include any kind of numerical/categorical data points in an application.

The HL7 FHIR standard does not restrict resources enough to remove all ambiguity in implementation details, as there are often multiple ways to create a FHIR resource with the same semantic meaning. This is partly due to the extension system in FHIR, which is presented as a way for developers to include

data which does not fit any of the existing resource categories on the FHIR server. This feature is intended to let developers tailor the specification to a given situation, as the base resources does not aim to cover all use cases. In practice, this means that developers have to coordinate among themselves, in order to determine the necessary restrictions for their use case.

Mental health data is, in the same way as any other type of medical data, sensitive and subject to security regulations. These problems can be mitigated during development by generating a data set of patients with values corresponding to what could be expected. Using the FHIR standard as an interface towards the data when generating the resources ensures the application supports interoperability towards future data sources. At any later stage, we can plug in a real data set in the FHIR format, and expect the technical aspects of the application to be compatible. This type of data set is available on open test servers [52], but is currently lacking in the mental health domain. The test servers also reset periodically, which makes them inconvenient to work towards. Due to these issues, the HAPI JPA server [53] was chosen for development of the artifact in this thesis. Sections 4.2.4 and 4.4 further describes how the resource server is populated with generated mental health data.

When making search queries, the resource server will return the results as a Bundle resource in the response. A Bundle is a composite FHIR resource, containing the results matching the search. Unpacking the Bundle will yield the resources in the appropriate formats. When a large number of resources are matching a search, the resource server can choose to return a subset of the matching resources in a series of pages [54]. Each response containing a page will also contain the link to the next page, if it exists. Retrieving multiple pages needs to be done sequentially as a result of this. Intuitively, this increases the time required to retrieve a collection of resources by  $numpages - 1$  Round Trip Times (RTT). The performance impact of this is discussed further in section 5.2.

The FHIR standard does not define security related functionality [55]. In practice, security has to be built on other protocols. The SMART on FHIR stack [56], using OAuth and OpenID Connect, offers potential as the default security mechanism. As SMART on FHIR is currently still in development, it was not used for development of security in the artifact.

## 2.5 Related Work

As mentioned in section 1.3.3, we found no clinical dashboard applications for use in iCBT, nor Internet-based therapy in general, through searching Google Scholar and Medline. While there is limited research on visualization dashboards within mental health, this is a well used technique in other disciplines of medicine. There are overlaps in information requirements between ICU and iCBT with with respect to the single patient. Ideas from the Intensive Care Unit (ICU) in hospitals can as such be adapted to the mental health setting, as visualizing patient data is a common goal.

When searching the scientific literature databases, multiple systems for visualization were identified. The studies that presented visualization methods which could be adopted to visualize the data in this thesis is included as related work. This section will present studies implementing medical data visualizations, and how these systems relate to the artifact developed in this thesis. The first three examples show visualizations of multivariate medical data for a single patient. All of the applications presented in these studies differ significantly from the design science artifact presented in this thesis, but contains one or two main similarities. The last study presents a dashboard for overview of a group of patients with diabetes.

### **2.5.1 Practical Visualization of Multivariate Time Series Data in a Neonatal ICU**

Lehmann et. al. [57] presents a clinical dashboard that was developed with the purpose of informing nurses in a neonatal ICU setting. The motivation for designing this dashboard was to enable providers to identify changes in the patient's condition in order to perform faster diagnoses. The study focuses on visualizing multivariate time series data. This is achieved by using star/spider plots, where each axis represents one measured variable for a patient.

The clinical dashboard displayed physiological measurements for a patient's heart, kidneys and lungs. Star plots were used to easier see how the variables change relative to other variables. Combined with animation representing patient state over time, this provided an overview of how the patients state changed during the measurement period. This idea has been adapted for visualizations in this thesis, with the main modification of applying user interaction instead of animation. Figure 2.4 shows an example of a star plot from Lehmann et. al. [57], paused during animation.

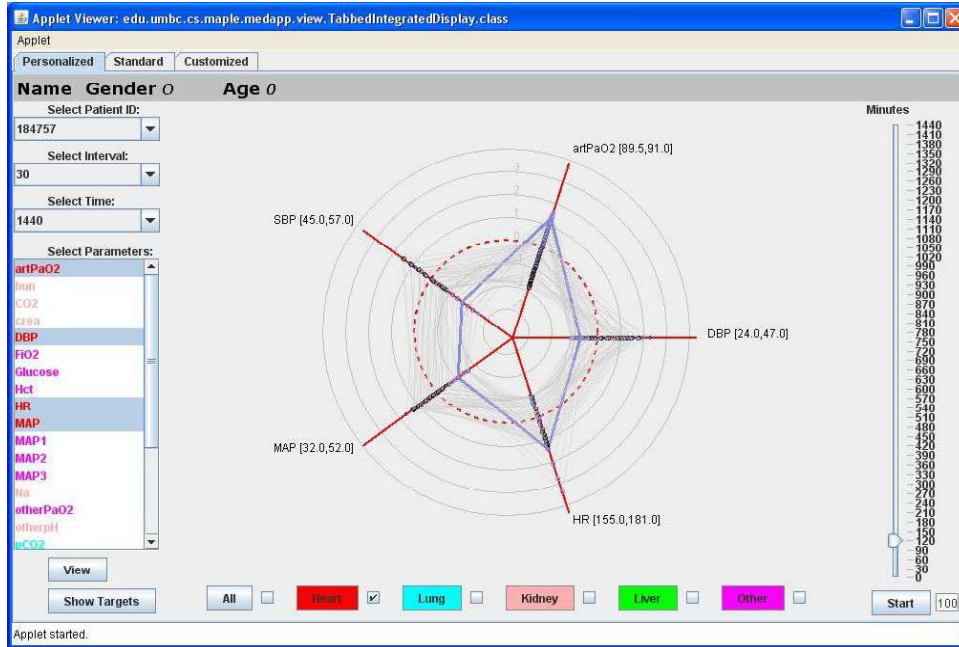


Figure 2.4: Animated spider chart paused, from Lehmann et. al. [57]

### 2.5.2 Multidisciplinary Intervention in Patients with Musculoskeletal Pain: a Randomized Clinical Trial

Brenbekken et. al. [58] presents a visualization tool named Interdisciplinary Structured Interview with a Visual Educational Tool (ISIVET). The goal of the study is to compare a multidisciplinary intervention (MI) performed with the tool, to a brief intervention (BI), for patients with musculoskeletal pain. Figure 2.5 shows one of the two figures used in ISIVET.

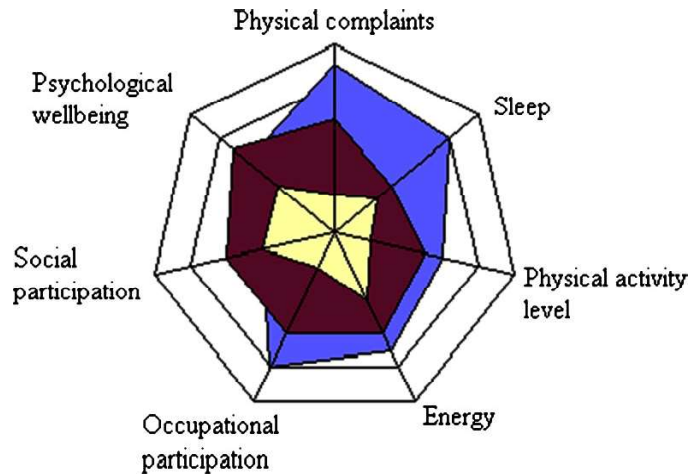


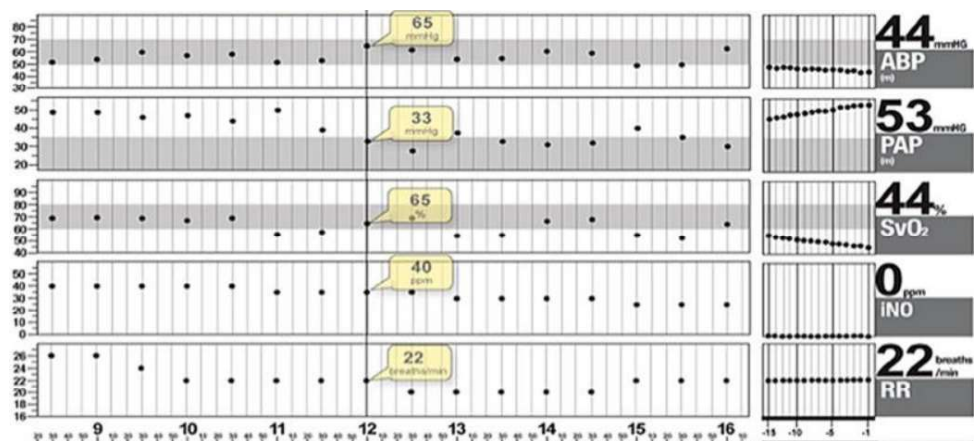
Figure 2.5: Spider chart in ISIVET

The spider chart in figure 2.5 displays a patients "quality of life". Each variable displayed is determined through cooperation between the patient and

the therapist. The three traces displayed in the chart represents "quality of life" filled in three times. This approach for visualizing patient reported data resembles the approach used for the spider chart in the artifact of this thesis.

### 2.5.3 Supporting Clinical Cognition: A Human-Centered Approach to a Novel ICU Information Visualization Dashboard

Faiola et. al. [59] presents a human-centered approach to creating a decision support tool for the ICU environment. The tool Medical Information Visualization Assistant (MIVA) was developed to visualize electronic medical record (EMR/EHR) data. Figure 2.6 shows "information visualization timelines" displaying multivariate biometric data.

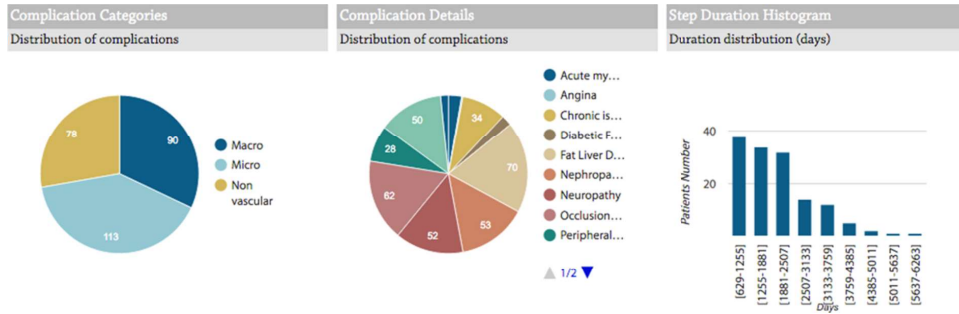


**Figure 2.6:** Temporal displays for biometric variables for a single patient in an ICU with MIVA 1.0

The approach for visualizing multivariate data in MIVA inspired the artifact in this thesis. Specifically, the line charts displaying single dimensional temporal values for each patient (see figure 3.4). The differences in the approach of MIVA and this thesis is substantial, but the main takeaway is stacking multiple charts of single dimensional time series.

### 2.5.4 A dashboard-based system for supporting diabetes care

Dagliati et. al. [60] presents a dashboard application to deliver Clinical Decision Support and outcome assessment for type two diabetes patients. Figure 2.7 shows some of the visualizations presented in an "outcome assessment and research support system (ORSS)" dashboard for diabetes patients. The application supports selecting sub-populations for further visualizations. Figure 2.7 shows the charts displayed for a selected set of patients.



**Figure 2.7:** Selected visualizations from dashboard for a patient population in Dagliati et. al. [60]

The pie chart on the left divides the patient population into a few high level categories. Similarly, the pie chart in the center displays a more detailed distribution of the selected patients. The right bar chart displays the distribution of the selected patient group with respect to the duration the patient had the matching level of disease evolution. All patients in this sub-population had the same level of disease evolution, and the bar chart shows each patients duration within this group. The ideas of using a pie chart to divide the patient group into high level categories, as well as using a bar chart to divide the patient group by temporal properties were adapted and used for the artifact in this thesis. Primarily the presentations of overview in the design science artifact was influenced by these ideas (see figure 3.1).

The CDS and visualizations presented by Dagliati et. al. considers a group of chronic patients. The data for these patients share some similarities with the patient data in iCBT, as the measurements span over a long period and considers outpatients. While the dashboard presented by Dagliati et. al. has multiple specialized features not discussed further here, the charts shown in figure 2.7 are generic enough to be adapted to the iCBT setting.

## 2.6 Section summary

Guided Internet-based Cognitive Behavioural Therapy is a mental health treatment form shown to have similar efficacy as standard Cognitive Behavioural Therapy while providing significant increase in efficiency. The patient-therapist contact in iCBT is mainly limited to asynchronous communication over the internet. As an effect of this, a therapist needs to rely on the available data to gain an overview of the state of her patients. The amount of data in iCBT is exceeding what a therapist can make use of in its raw format. One of the main ways to track the progress of patients in iCBT is through standardized questionnaires. In current practice at eMeistring, therapists have to consider each data point manually, through pdf files or charts displaying single questionnaires. Gaining an overview of one or multiple patients requires significant time, and is cumbersome for the therapist within the current workflow. Through the application of visual analytics and dashboard techniques, therapist workflow can be improved with respect to efficiency and reduction of cognitive load.

The patient's response to the exercises given in a treatment module is also important data for therapists to consider. In current practice, the exercises are mainly free-text based. Structuring this data for further analysis is an important and interesting topic for further research. Quantitative measurements in the form of standardized screening questionnaires and temporal numerical observations will be the focus for the artifact in this thesis.

The data from questionnaires can be visualized to show patient progress over time. By applying data visualization and aggregation techniques, a dashboard application can provide insight into both a therapist's group of patients, and each single patient. Adopting HL7 FHIR resources as the storage format of the filled questionnaires will enable the dashboard to be general enough to support iCBT treatment programmes for multiple mental health disorders. Using the FHIR standard will also be beneficial for integrating data from multiple systems into a dashboard application, as having a single format for the data will reduce the complexity of system interoperability. As there exists limited work on clinical dashboards for Internet-based mental health therapy, techniques for data visualization in other medical domains are adopted and implemented in the artifact.



## 3 Method and Design

This section will present the artifact. The process of designing and developing the artifact will be described first. Following this, an overview of the user interface along with examples of use cases is presented. The specifics of converting screening questionnaires to time series is then described. Finally, the visualization for multidimensional time series implemented in the artifact is presented.

### 3.1 Design process

Throughout the development of the artifact, meetings were held to evaluate the artifact design and to discuss its features. These meetings were held at roughly two to four week intervals. During these meetings, the artifact was presented, and changes as well as new features were discussed. The supervisors of this thesis, who have experience working on iCBT applications, were present at these meetings. Other master students in health informatics were also present. It would have been beneficial for the development of the artifact to have therapists attend these or similar meetings, as after all, they are the intended user group. Unfortunately, this was not possible. The iterative development process was carried out until the artifact satisfied the design science objectives. Technical details about the implementation throughout this process is presented in section 4.1.

The development of the artifact can roughly be divided into four iterations. The iterations had slight overlaps, but followed the sequence described below. During each iteration, we set goals for the development of the artifact, through discussion in the meetings described above. The artifact was evaluated informally at these meetings, and new goals were determined for the subsequent development period. The overarching goals of each iteration were the following:

1. Problem identification and initial ideas for solutions.
2. Identifying data types, adopting HL7 FHIR and generation of data.
3. Implementation of visualizations, and integrating them in a dashboard.
4. Scaling the application to support more data points.

**Iteration one:** The initial concerns of designing the artifact considered how the mental health screening data could be presented in a better way than through patient filled questionnaire forms. The patterns described in section 3.3 were identified here, and some methods for visualizing the data were proposed. The initially suggested charts were heat maps, line charts, and spider charts. We decided the heat maps were too verbose for practical use. The other charts were integrated in the GUI of the initial prototype, described in section 4.1.2.

**Iteration two:** The FHIR standard was then adopted, and mental health data generated as described in section 4.4. We discovered that supporting the unstructured exercise data of eMeistring (e.g. plans, activities, and follow-up on these) in its current form would be difficult. To keep the dashboard generic

enough to support multiple forms of iCBT, we chose to focus on utilizing the data we could express in the FHIR format.

**Iteration three:** The overview page, with the corresponding visualizations were designed next. Simple visualizations were chosen for this view, as it is intended to give fast overview rather than provide deep analysis. The interactive spider chart was also designed to visualize the QuestionnaireResponse resources for each patient. This iteration also considered consolidating all the visualizations into a functioning dashboard.

**Iteration four:** In this iteration, we scaled the dashboard solution to support additional patients per therapist, through the generation of more data. This process uncovered some performance problems, where the main bottleneck was tracked down to be interaction with the HAPI JPA server (see section 4.6.1). Most of the remaining development time was spent mitigating these issues, as described in section 4.6.1

## 3.2 Artifact description

The artifact is a web-app consisting of two primary views. Throughout this thesis, we will refer to these as the Master View and Detail View. The purpose of the artifact is to display the status of both a group of patients as well as each single patient. The Master View presents an overview of all the patients, and the Detail View presents visualizations of the data for a specific patient. The application uses FHIR to interface towards the data source, and does as such have the potential to support a wide range of iCBT treatments.

### 3.2.1 Master View

The Master View is the landing page of the application. Here, a therapist is presented with an overview of her group of patients. The view is designed to comply with Shneiderman’s mantra [48], giving an overview first with possibilities to filter, and then details on demand. Status variables are presented to describe each patient. For particularly urgent patients, a warning string is presented. Each patient will have a progression string and numerical urgency score (if the data is present), with optional flags and warning strings. The calculation of these variables are discussed further in section 4.5.1. Figure 3.1 shows the entire Master View. The numbers are overlaid for identification of the GUI elements in the following paragraph.

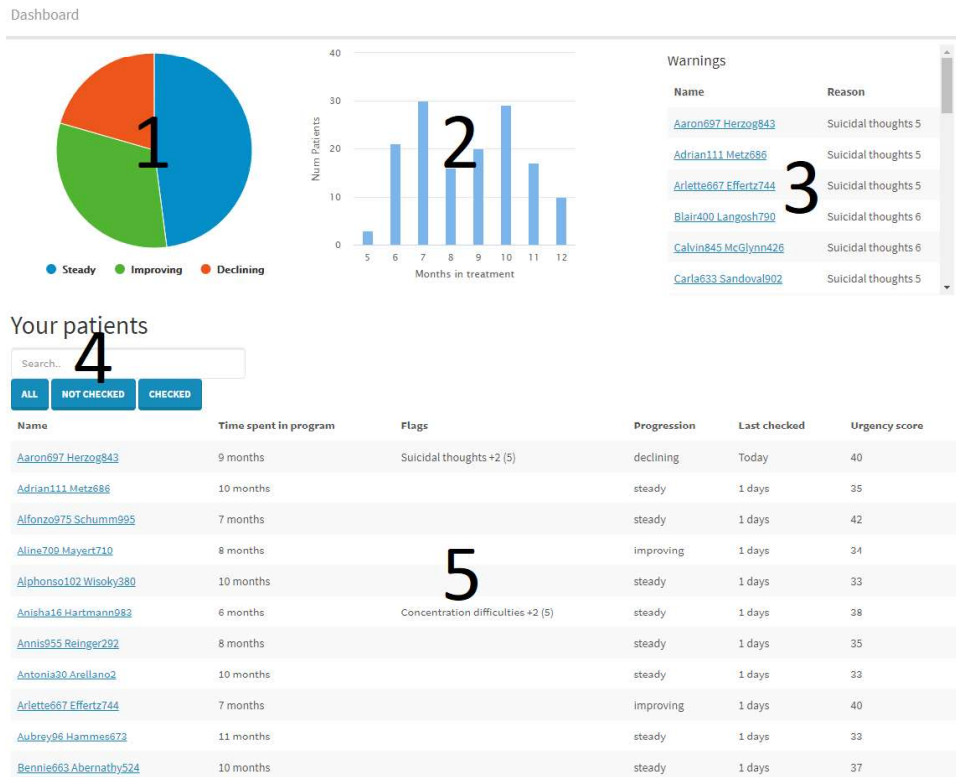


Figure 3.1: Master View

The main GUI element in the Master View is the table over all the patients (5). The status variables described in the previous paragraph are presented for each patient. This table is sortable on any column, and can be filtered based on a search string, or if they have already been checked by the therapist this week (4). Patients with urgent needs are presented in the warning section (3). The pie chart represents the state of the patient population in terms of the progression variables (1). The bar chart groups patients based on how long they have been in the treatment program (2). All together, this elements in this view presents the therapist with a fast overview of the state of her group of patients. Figure 3.2 and 3.3 show GUI element one through four presented here.

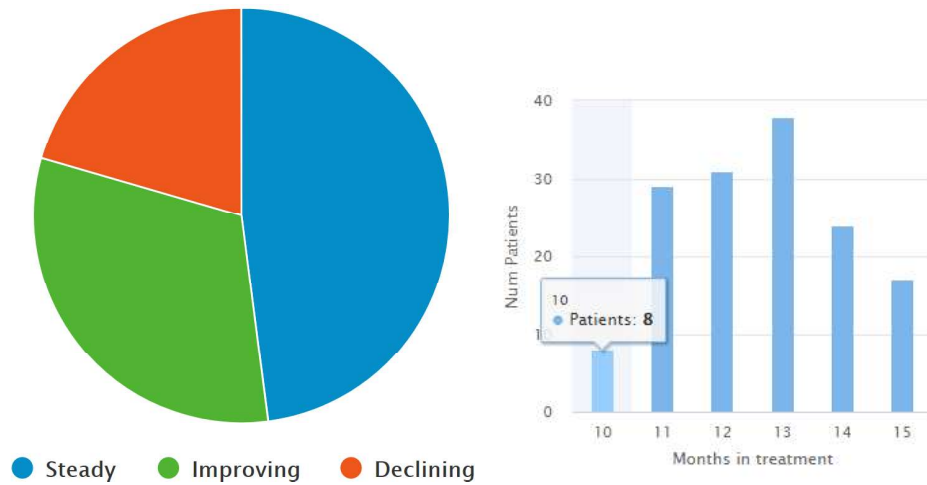


Figure 3.2: Pie- and Bar chart in Master View (1 and 2)

The pie chart provides an overview of the group of patients based on the progression variables. Progression variables can take the values "improving", "steady", "declining", as described in section 4.5.1. As the progression strings are known, color coding is applied to the pie chart. Intuitive color choices have been used here, with green representing improvement and red representing decline. Clicking any of the slices of the pie chart will place the patients with the matching slice on top of the table. The bar chart shows how long patients have been in the treatment.

### Warnings

Name	Reason
<a href="#">Aaron697 Herzog843</a>	Suicidal thoughts 4
<a href="#">Adrian111 Metz686</a>	Suicidal thoughts 5
<a href="#">Alfonzo975 Schumm995</a>	Suicidal thoughts 4
<a href="#">Aline709 Mayert710</a>	Suicidal thoughts 4
<a href="#">Alphonso102 Wisoky380</a>	Suicidal thoughts 4
<a href="#">Anisha16 Hartmann983</a>	Suicidal thoughts 4

### Your patients

Search..

ALL NOT CHECKED CHECKED

Figure 3.3: Warnings (3), and tools for search and filtering of table (4)

Combining the table with the filters seen in figure 3.3, two main options for workflow are presented. Therapists who want to check all their patients every week (or another set interval) has the option to filter out patients they have already seen. The other option is treating the patients based on the overview information. The therapist can go through the table row by row, and check each patient if they want to. Alternatively, methods to sort and filter the table is

presented to make it easier for the therapist to find the information they require, for how to organize their work session for the day. The warning function used in this example checks all patients latest MADRS QuestionnaireResponse, and displays the patients with a score  $\geq 4$  on the suicidal category. When using the MADRS questionnaire, patients answering  $\geq 4$  on this category should be considered for admission [61]. When using other questionnaires for the treatment, the calculation of warning variables can be adjusted.

### 3.2.2 Detail View

Clicking on a patient’s name will direct the browser to a detailed view of the patient. The data presented in this view, as well as the rest of the application, is generated as described in section 4.4. For the generation of the patient resources, the Syntea patient generator [62] library was used. The background information displayed in figure 3.4 and 3.5 is the data on the FHIR Patient resource.

The available visualizations in this view is dependent on the data available from the resource server. All QuestionnaireResponse and Observation resources for the patient is pulled from the FHIR server. Any QuestionnaireResponse resources is visualized by default. The therapist can then select which of the available Observation resources should be shown. The Observation resources can be visualized as regular line charts. Figure 3.4 shows the entire Detail View, with numbers overlaid similarly to the Master View in figure 3.1.

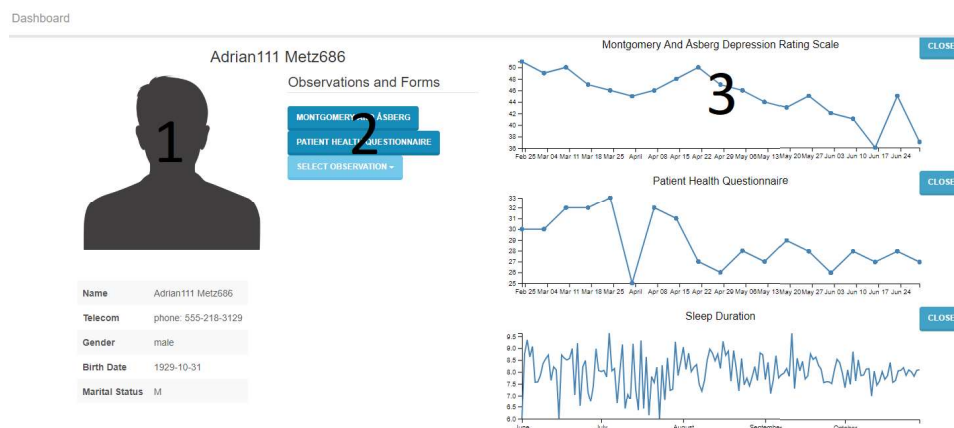
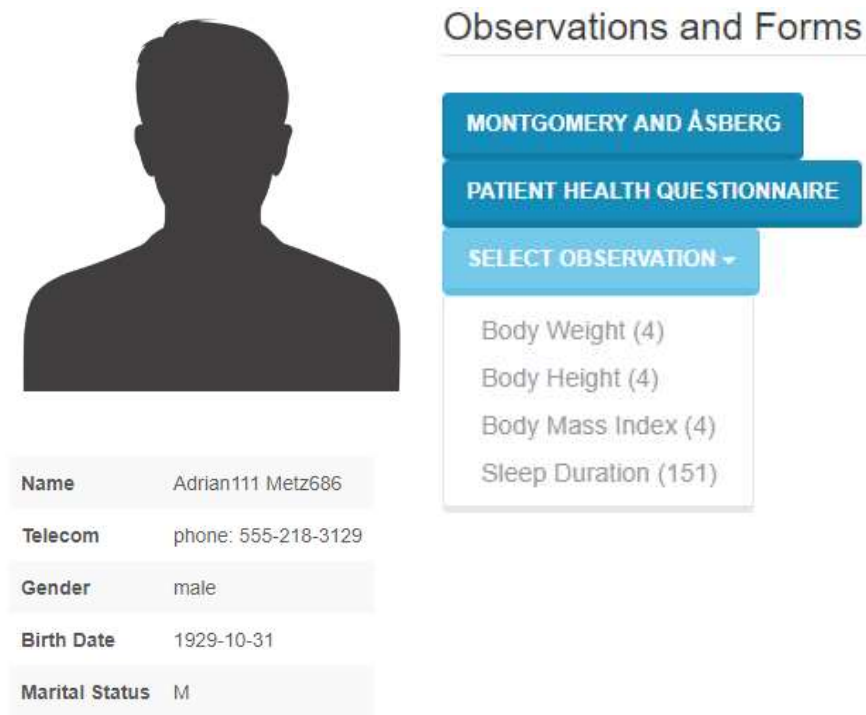


Figure 3.4: Detail View

Visualization of the QuestionnaireResponse resources requires a multidimensional approach. The method used in the artifact is an interactive combination of Spider Chart and line chart, as shown in section 3.4. These charts are displayed by clicking the appropriate button in (2). As the sum of the answers in psychometric screening questionnaires is informative, line charts displaying the sum over all answers in each QuestionnaireResponse resource is displayed by default. This chart presents an aggregate overview of the data, where further details can be discovered in the spider chart. The interactive spider chart is described further in section 3.4.

The therapist can customize this view by adding or removing any line charts, given the available data. For any QuestionnaireResponse resources with at least two data points, a line chart for this questionnaire will be shown as default. The line charts will be persisted between user sessions. This is intended to help integrate the application with user workflow, and reduce the number of actions the therapist needs to perform. As which data types are relevant varies between patients, persisting the state of the view can also help the therapist gain a faster overview of a patient. As described by Nielsen, recognition is preferable to recall [63][64]. By being presented with the state of the view from the previous work session, the user will recognize the patient, and remember the patient’s problems faster.

The visualizations can display a wide range of data types. Because of this, the interpretation of the charts will be dependent on which data is being presented. For the questionnaires used in the examples presented in this thesis, lower scores are better. A downwards trend in the charts can as such be considered a positive development. Data types where higher values are better than lower values would need to be interpreted differently. With regards to automatic interpretation and decision support, this is further explored in section 4.5.1. Figure 3.5 shows the patient’s background information, and how observations can be selected.

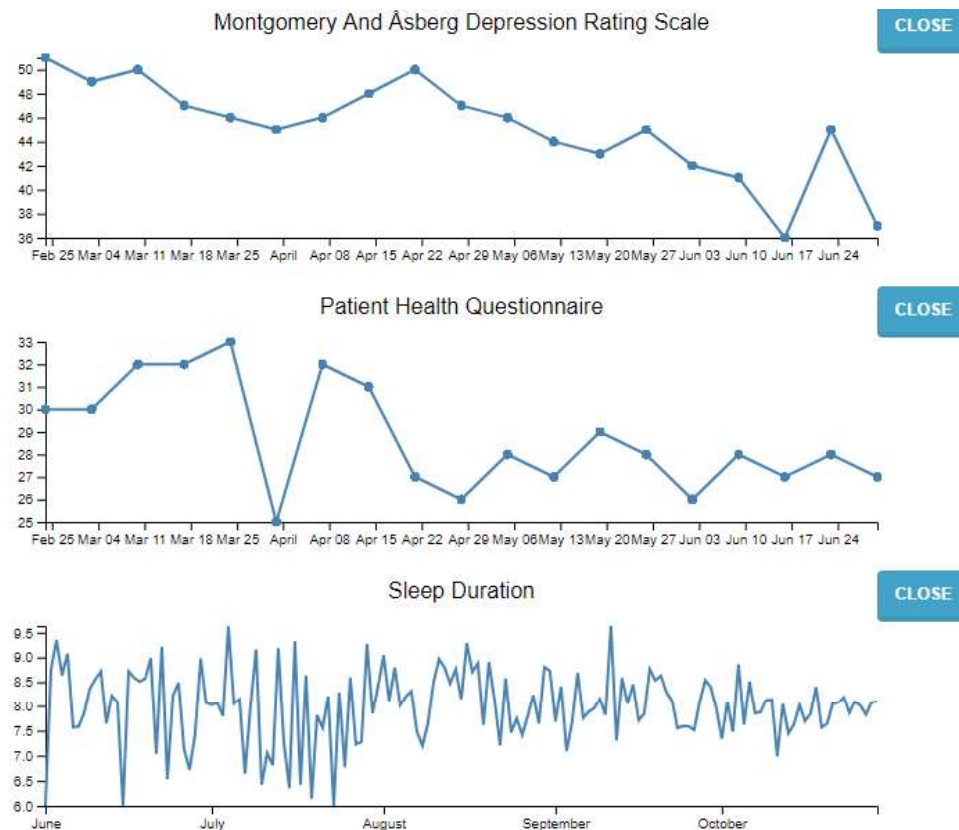


**Figure 3.5:** Picture and background information (1) and selectors (2)

The background information in figure 3.5 displays the data from the Patient resource. In this case, the fields in the Patient resource generated by Synthea is shown. What kind of background information is appropriate will depend on the situation in practice. The fields displayed here can be changed by extending

the Patient resource. As the Patient resource is filtered before it is displayed in the artifact, changing the displayed fields would require a small implementation change.

In figure 3.5, the top two buttons will display interactive spider charts when selected. The "Select Observation" drop-down menu presents a list of all available Observation resources for the patient. The number of data points are shown in parenthesis. Only Observations with at least two data points, and the "valueQuantity" field are displayed. In the example shown in figure 3.5, the "Body Weight", "Body Height" and "Body Mass Index" Observations are generated by the Synthea library, described in section 4.2.4. The "Sleep Duration" Observations are generated by a Python script described in section 4.4. Figure 3.6 shows examples of line charts displaying data from QuestionnaireResponse and Observation resources.



**Figure 3.6:** Line charts (3)

The top two line charts in figure 3.6 show the progression of the patient, represented by the questionnaires MADRS and PHQ-9. Each data point is the sum of the answers in the QuestionnaireResponse at a given time. Why the sum is useful to display is explained further in section 3.3. The axes are normalized to the range between the maximum and minimum values of the data points. For these specific examples, a downwards trend (symptom improvement) can

be observed in the MADRS line chart, and stabilization could be observed in the PHQ-9 line chart.

The chart labeled "Sleep Duration" shows how many hours a night the patient slept. This chart could show any of the measurements displayed in figure 3.5. This generated example shows how higher frequency data can be displayed in the artifact. If there are less than 80 data points, the blue circles on the traces are hidden to improve the readability of the chart. A reduction in volatility over the measured interval can be observed on this chart. For a real patient, this could indicate improvement in sleeping habits

### 3.3 Psychometric screening data as time series

The questionnaires given to patients can take a form similar to the example given for the MADRS questionnaire in figure 2.1. For some questionnaires, both patient and therapist can fill in the answers. Other forms are specialized towards either the patient or therapists. As an example: The first question in the MADRS form should be answered by the therapist. This question is removed in the MADRS-S version. Other variations in the MADRS questionnaire considers some versions only accepting even numbers, while other versions accepts all numbers from zero to six. The sum over the answers in MADRS can be compared to certain thresholds in order to determine the severity of the depression [61].

As described above, multiple versions of the same mental health screening questionnaire can be used. Due to this, there is a wide range of unique questionnaires suitable for iCBT. The questionnaires does, however, mainly conform to the following patterns:

- Each question has a numerical answer, or an ordinal answer which can be converted to a number.
- The numbers of the answers are in the same range for all questions in the questionnaire.
- The set of answers can be aggregated to an overall evaluation, for example by summarizing all the answers in the questionnaire. This aggregated result can in certain cases be compared to threshold values in order to determine a diagnosis.

All questionnaires mentioned in section 2.1 follows these patterns. As the questionnaires are filled out multiple times over an interval, the set of answers can be represented as unevenly spaced multidimensional time series. Some applications of time series analysis requires constant intervals between the measurements. In this case the intervals can not be assumed to be constant, as patients could be delayed or inconsistent in delivering the questionnaires. For visualization purposes unevenly spaced data points are acceptable. We show one method for visualizing this type of data in section 3.4.



Considering the MADRS questionnaire, table 3.1 shows a more compact representation of the answers. Further compression of the data is possible, by only saving the vector of answers. Given the questionnaire, the indices in the vector can map each answer to the correct question.

1. Reported sadness	4
2. Inner tension	4
3. Reduced sleep	5
4. Reduced appetite	6
5. Concentration difficulties	4
6. Lassitude	3
7. Inability to feel	4
8. Pessimistic thoughts	4
9. Suicidal thoughts	2

**Table 3.1:** Answers to the MADRS questionnaire in tabular format

	$t_0$	$t_1$	$t_2$	$t_3$
1. Reported sadness	4	6	4	4
2. Inner tension	4	2	4	6
3. Reduced sleep	5	2	3	4
4. Reduced appetite	4	2	2	0
5. Concentration difficulties	2	5	2	0
6. Lassitude	4	2	3	4
7. Inability to feel	1	2	4	2
8. Pessimistic thoughts	4	0	3	4
9. Suicidal thoughts	0	2	3	0

**Table 3.2:** Multidimensional time series example

Formatting the answer vectors as multidimensional time series can produce the frame shown in table 3.2. This format is useful for reasoning about further processing of the data. In the artifact, the data is not represented in this exact format, but as a similar dictionary representation. The following processing techniques can create one dimensional vectors, suitable for visualization:

- **Select row:** The resulting vector will contain a time series for the category/question of that row. Line chart is a suitable visualization for this time series.
- **Select column:** The resulting vector represents a questionnaire at one point in time. Bar charts or spider charts are suitable visualizations for this vector of ordinal values.
- **Sum over columns:** The resulting vector will contain a time series with the sum of categories at each time step. Line chart is a suitable visualization for this time series. The sum of categories can be used to determine treatment options in questionnaires like MADRS. For example: A sum  $\geq 36$  is considered very severe depression [61].

The resource server stores each set of answers as FHIR QuestionnaireResponse resources. Figure 3.7 shows an example of how the MADRS questionnaire can be represented as a FHIR QuestionnaireResponse resource (clipped after two questions).

```
1  "resource": {
2    "resourceType": "QuestionnaireResponse",
3    "id": "75289",
4    "meta": {
5      "versionId": "1",
6      "lastUpdated": "2019-03-25T13:54:32.158+01:00"
7    },
8    "questionnaire": {
9      "reference": "Questionnaire/42220"
10   },
11   "status": "completed",
12   "subject": {
13     "reference": "Patient/2"
14   },
15   "authored": "2018-03-18",
16   "item": [
17     {
18       "linkId": "0",
19       "text": "Reported sadness",
20       "answer": [
21         {
22           "valueInteger": 4
23         }
24       ]
25     },
26     {
27       "linkId": "1",
28       "text": "Inner tension",
29       "answer": [
30         {
31           "valueInteger": 4
32         }
33       ]
34     }
35     ...
36   ]
37 }
```

**Figure 3.7:** Example of a HL7 FHIR QuestionnaireResponse resource for the MADRS questionnaire

This example could potentially contain more information, as some optional fields have been left unfilled. The QuestionnaireResponse resource is in this case a verbose representation of table 3.1. The linkId fields in the QuestionnaireResponse maps to the linkId fields in the corresponding Questionnaire resource. These ids can be used in conjunction with the resources to recreate question-

answer sets. As multiple QuestionnaireResponse resources can have the same Questionnaire, this division of questions and answers can reduce the amount of text contained in each QuestionnaireResponse. When requesting large amounts of QuestionnaireResponse resources this can provide performance benefits.

To summarize, the storage format is not necessarily the best format for reasoning about how to process the data. Using the FHIR format as the primary data representation, however, is beneficial for system interoperability. Converting the resources into a more compact form after requesting them can be necessary for visualization.

### 3.4 Visualizing multidimensional time series

With the goal of visualizing data from the format shown in table 3.2, we built a custom visualization, shown in figure 3.8. As a high degree of customizability was required, we chose the D3.js library for this task. A spider chart was chosen to represent a set of answers at one point in time. As the dimensionality of the intended data is mid-range, spider charts are a suitable option for visualizing the questionnaire data. For mid-range, we consider roughly between four and 20 dimensions. Three dimensions or less could be visualized in a line- or bar chart, and over 20 would require more complex techniques. The spider chart contains two traces, one representing the most recent QuestionnaireResponse and the other trace representing an earlier QuestionnaireResponse selected by the user. Using the data representation in table 3.2, the blue trace represents column  $t_{n-1}$  where  $n$  is the number of data points. Similarly, the orange trace represents any column in  $t_0$  to  $t_{n-2}$ . Two traces were chosen to easier display changes in patient state, where blue being lower than orange indicates improvement along an axis.

A mechanism to enable the user to select the date of the orange trace was designed next. To implement this, we used a line chart summarizing the values of all axes of the spider chart at each time step. When the therapist hovers this line chart, the orange trace of the spider chart updates to the date closest to the cursor. The orange vertical line in figure 3.8 represents the selected date. This approach allows the therapist to quickly compare patient progress along each axis in the spider chart. Furthermore, each axis can be clicked by the user to display a line chart over the time series in the selected category. This enables the therapist to customize which visualizations to see for each patient. Figure 3.8 shows the interactive spider chart visualizing mental health screening data.



**Figure 3.8:** Interactive Spider Chart

The visualization in figure 3.8 is designed for any series of QuestionnaireResponse resources, given the format previously described and at least two data points. For the visualization to be correct, the QuestionnaireResponse resources presented should also reference the same questionnaire. This issue is handled when retrieving the resources from the FHIR server. The artifact will group the QuestionnaireResponse resources based on their referenced questionnaires. Buttons will be rendered in the GUI to select interactive spider charts for any of these groups with enough data points (at least two resources). As the artifact treats the resources in a generic manner, no validation is performed towards the semantic content of the resources. Because of this, the quality of the visualizations depends on the data available from the resource server. The spider chart portion of the visualization is based on [65], but heavily modified.

## 4 Implementation

While the previous chapter presented the artifacts design, this chapter provides a more in-depth explanation of the implementation. More technical descriptions are provided on how calculations are carried out. First, the tools used for development of the artifact is presented. The system's architecture and update patterns are then described. Generation of mental health data follows. The status variables, how they are calculated, and potential for further extension is then presented. The scalability of the artifact and further potential for optimization is then described. Finally, we discuss potential security considerations.

### 4.1 Tools and frameworks used during development

This section and section 4.2 presents the main tools used to create the artifact. Only the main technologies used in the artifact are presented. Tools like IDEs and version control were also used, but the specifics of these tools are not necessary to include. The first two tools presented in this section were used for iteration one, described in section 3.1. The SMART on FHIR JavaScript client was used for iteration two and three. The SMART on FHIR Python client was used for iteration two and onwards.

#### 4.1.1 Matplotlib and Seaborn

We created the initial prototypes for the visualizations with the Python visualization libraries Matplotlib [66] and Seaborn [67]. These tools allow for rapid experimentation, as they provide many pre-packaged visualization options. Both libraries are designed to create locally running visualizations for a local data set. The libraries present good solutions for fast overview of a data set, but they do not present web-based visualizations.

#### 4.1.2 Dash

After discovering the initial options for visualizations, we designed an interactive prototype for further testing. We used the Dash framework [68] to create this prototype. Dash provides a lightweight web server made for fast creation of charts by only using Python. Due to the common language of Python, Dash is partially compatible with Matplotlib and Seaborn. Adapting the Matplotlib visualizations towards Dash was as such unproblematic. This was useful in order to get a fast GUI prototype which could be used for discussion with the therapists. The prototype was presented in the meeting at the eMeistring clinic, described in section 1.3.2. After a short period of using Dash, we found the framework to not be flexible enough to support all our requirements. We then considered the JavaScript visualization libraries D3 and Highcharts, described in section 4.2.5. As Dash did not provide sufficient flexibility, we decided switch to the heavier web server framework of ASP.NET MVC, described in section 4.2.1.

#### 4.1.3 SMART on FHIR JavaScript client

After adopting the FHIR standard, as described in the previous section, we made more changes to the tools used. For the handling of FHIR resources,

we initially used the SMART on FHIR JavaScript client. SMART on FHIR is designed to provide a full security tech stack with OAuth 2 and OpenID Connect. As the design science artifact does not implement security, only the RESTful operations of the client were used. Due to the performance issues discussed in section 4.6, we moved the calls to the FHIR resource server to the backend. As such, the SMART on FHIR JavaScript client was replaced by the C# FHIR client [69], which is included in the final artifact.

#### **4.1.4 SMART on FHIR Python client**

The SMART on FHIR Python client library [70] was used as part of the data generation process to convert raw data into QuestionnaireResponse and Observation resources. The library also provided the necessary functionality for CRUD interactions with the resource server, allowing for directly uploading the resources after they were generated.

## **4.2 Tools and frameworks in the final artifact**

Most of the tools presented in this section were integrated in the artifact in the second iteration of the development, described in section 3.1. The exceptions are the JavaScript visualization frameworks Highcharts and D3 which were used from iteration three.

### **4.2.1 ASP MVC**

The framework used for the web server in the artifact was ASP MVC 5 .Net. The server was mainly used to serve the data to the views initially, with calculations being performed on the client. We chose ASP MVC initially because it has the potential to integrate seamlessly with other powerful .NET frameworks, for example ASP.NET Identity [71] and Entity Framework [72]. We did, however, decide not to use either of these frameworks for the final prototype as they are designed towards relational databases. As described in section 4.2.2, we decided to use the document-based database MongoDB. The official C# FHIR client library [69] was used to make requests between the application server and the resource server.

### **4.2.2 MongoDB**

The web server uses MongoDB as the database, along with the official MongoDB C# driver [73]. We chose MongoDB due to its compatibility with the JSON format. MongoDB stores data in collections of BSON (Binary JSON [74]) documents. When requesting resources from the FHIR server, the response is sent as JSON format (could optionally have been XML). This format is also used for JavaScript on the client, as JSON can be directly casted to JavaScript objects. These interactions entails that reading objects from the database and then passing them to the client does not require a lot of code.

### **4.2.3 HAPI JPA**

The HAPI JPA [53] Java reference implementation of FHIR was used as the resource server. The server was only modified to disable paging of resources,

which is further discussed in sections 4.6 and 5.2. The artifact also supports resource servers with paging, although we do not recommend this due to the findings described in 5.2. To ensure compatibility with other FHIR servers, no other modifications were made. FHIR version DSTU3 was used for this server, as version R4 was released during the development of the artifact.

#### 4.2.4 Synthea

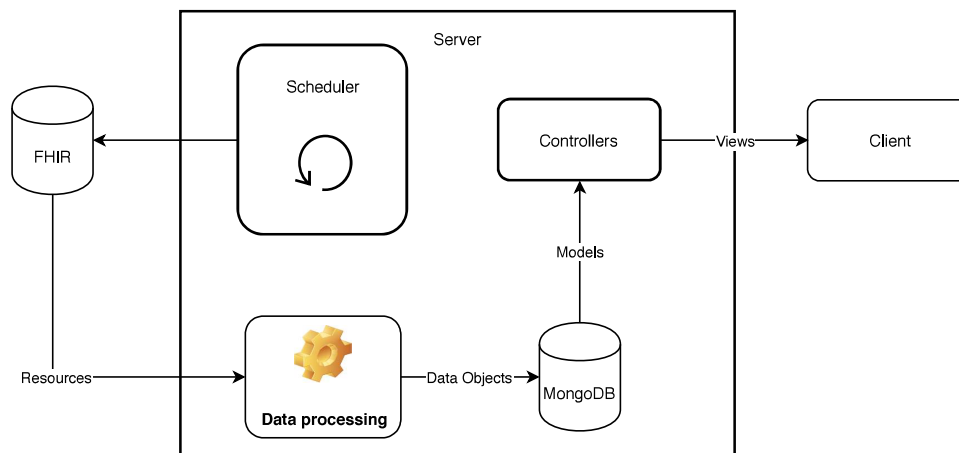
The Synthea Patient Generator [62] generates realistic synthetic patient data in FHIR format. For the development of the artifact, we used Synthea to generate 147 synthetic patients in JSON format. We then uploaded these resources to the HAPI server. The generated data did not contain all the required resources for the artifact, particularly within mental health. For the mental health data, we generated the remaining resources as described in section 4.4.

#### 4.2.5 Highcharts and D3

Highcharts [75] and D3 [76] are both JavaScript libraries for data visualization. Highcharts provides customizable pre-made solutions built on D3, and is a commercial project which is free for non commercial use. We used the Highcharts library to create the pie- and bar chart in the Master View. D3, or Data Driven Documents, is a library with with a high level of customizability. We used D3 to create the interactive spider chart described in section 3.4.

### 4.3 System Architecture

The architecture of the artifact is mainly based on the Model View Controller (MVC) pattern. As the models require data from an external resource server, the complexity of the architecture increases. Figure 4.1 shows the flow of data between the high level components in the artifact. This flow is further explained in section 4.2.



**Figure 4.1:** Architecture model

The "Data processing" step in figure 4.1 only considers processing of patient data. More specifically, this processing considers the logic of calculating the

status variables, wrangling data, and deserialization/serialization of the FHIR resources. Processing the therapist specific fields of the model requires no interaction with the FHIR server. As this interaction is the main performance bottleneck (see section 5.2), actions related to therapist specific data is fast enough to be processed on each request. The FHIR server in figure 4.1 is an implementation of the HAPI JPA server. This server is replaceable by any resource server implementation conforming to FHIR version DSTU3.

The client component of figure 4.1 represents the web browser of the therapist. All data visualizations are processed and displayed here. The data is already mostly processed when it is delivered to the client through the model. As such, the logic of the visualizations are completely decoupled from the backend/server.

The scheduler and resource server can run in the same virtual environment as the application server, but this is not a necessity. For example if the artifact is deployed to a cloud environment, a cloud scheduler can be used. Running all component on the same virtual machine is most likely the simplest solution for small scale projects.

The scheduled update process runs independently of user interaction. When the client requests a view, the data is already available. As a result of this, the data will not be available in real time. How recent the displayed data is will depend on update frequency.

### 4.3.1 Update flow

Figure 4.2 presents a more detailed description of the data processing module of figure 4.1. QuestionnaireResponse resources are abbreviated to "QRs" in the diagram. The sequence is initiated by the scheduler in figure 4.1. The application server will first fetch the Patient resources and their corresponding QuestionnaireResponse resources from the FHIR server. These resources are processed to create the model for the Master View. The QuestionnaireResponse resources are then stored in the database.

As the FHIR specification prescribes a reference in the questionnaire field of QuestionnaireResponse resources, some additional calls needs to be made to retrieve the necessary Questionnaire resources. In order to get the correct questionnaire name/title for a given QuestionnaireResponse, a dictionary between the URIs and questionnaire names are made during the update process, for each patient. This is used in the models for the Detail View, for displaying the correct names of the questionnaires in the view.

The reason for storing the QuestionnaireResponse resources is to decouple the processing of the models from the Detail Views to the Master View. Storing and retrieving the models in the same process incurs a slight increase in processing time, but reduces code complexity. For development and debugging, being able to update one type of model at a time allows for rapid development.



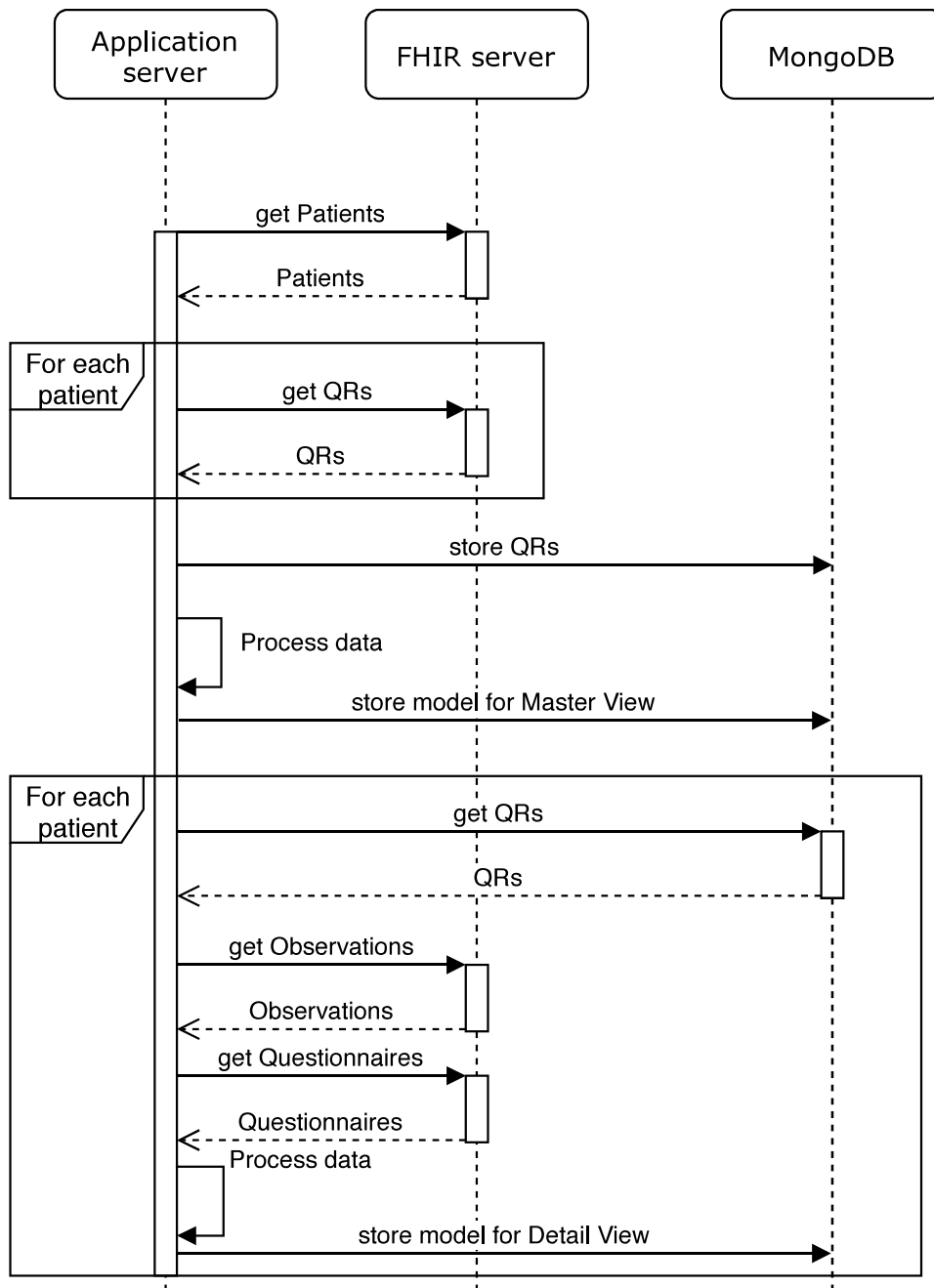


Figure 4.2: Sequence diagram for the update flow

After processing the Master View models, each patients Detail View model is created. The QuestionnaireResponse resources are retrieved from MongoDB, and the Observations from the resource server. All unique references to Questionnaires in the QuestionnaireResponse resources are also retrieved. These resources are serialized and stored as fields in the models for the Detail View. The models are then persisted in MongoDB.

## 4.4 Mental health data generation

After generating the synthetic patients with Synthea, the next goal was to simulate realistic mental health data. QuestionnaireResponse, Questionnaire and Observation resources were generated. The QuestionnaireResponse described in section 3.3 is an example of one of these resources. To simulate patients answering a questionnaire over a time period, the resources for the MADRS questionnaire was generated first. To achieve this, we created a Python script to populate the FHIR server with resources. In this script, the data is first generated as multidimensional time series. Next, the data is converted to FHIR resources using the SMART on FHIR Python client [70]. Finally the resources are uploaded to the FHIR server. We used the following Python code to generate realistic looking data, by simulating variance and trend:

```
default_max_timesteps = 30
for i in range(len(dates)):
    decrement_fraction = 1 / max(default_max_timesteps, len(dates))
    mu = max_val - (min_val + 2 * (i * decrement_fraction))
    variance = 1 - (i * decrement_fraction)
    sigma = math.sqrt(variance)
    col = np.random.normal(mu, sigma, 9)

    #clamp values: min <= x <= max
    col = [min(max(x, min_val), max_val) for x in col]
```

**Figure 4.3:** Code snippet for simulating realistic answers to questionnaires

The "col" variable is an array containing the generated answer vector at a given time. The answers are drawn from a normal distribution given the  $\mu$  and  $\sigma$  at that time step. The decreasing mean and variance is intended to indicate an improving trend for the patient. This trend can be recognized in the visualizations of the artifact. The generated dates are evenly spaced dates in a semi-random interval. The interval is chosen as the interval between a random date in the last seven months of 2018, to a random date in the first five months of 2019.

We repeated this process for all patient resources on the FHIR server, for the MADRS-S and PHQ-9 questionnaires. To generate sleep data, we also created a similar script to create Observation resources. The main modification was to generate single values instead of an array, and having constant mean over the time series. An example of a single sleep observation is displayed in figure 4.4. The main fields used in these Observation resources by the artifact is "valueQuantity.value" and "effectiveDateTime". Almost all fields displayed here are required by FHIR DSTU3.

Due to the nature of the data, we can not expect the generated data to show similar properties to real data. It would, for example, be reasonable to expect certain categories to be lower than other categories on average. In MADRS, the suicidal category is more serious than the other categories, which could

```

1  "resource": {
2    "resourceType": "Observation",
3    "id": "57473",
4    "meta": {
5      "versionId": "1",
6      "lastUpdated": "2019-03-04T17:05:53.662+01:00"
7    },
8    "status": "registered",
9    "code": {
10     "coding": [
11       {
12         "system": "SNOMED CT",
13         "code": "248263006",
14         "display": "Sleep Duration"
15       }
16     ]
17   },
18   "subject": {
19     "reference": "Patient/2"
20   },
21   "effectiveDateTime": "2018-06-03",
22   "valueQuantity": {
23     "value": 6,
24     "unit": "h",
25     "system": "http://unitsofmeasure.org",
26     "code": "h"
27   }
28 }

```

**Figure 4.4:** Example of a HL7 FHIR Observation resource measuring a patients sleep duration for one night

cause the average score for this category to be lower. We could also expect different patients to have different problems, represented by differing answering patterns. While none of these properties were accounted for in the generated data, some trend and change in variance is beneficial for demonstrating how these phenomena would impact the visualizations.

## 4.5 Status variables

The artifact presents both visual elements and status variables to provide the therapists with decision support. Section 3 presented the visualizations. This section will present how the status variables for decision support are calculated.

Status variables are calculated for each patient in the Master View. The variables are represented as strings, and are calculated based on the content of the patients QuestionnaireResponse resources. The variables are expressed as

”Time spent in program”, ”Progression”, ”Flags”, ”Warnings” and ”Urgency score” in the GUI of the artifact. A column for ”Last checked” is also presented for each patient, describing when the therapist last looked at the patient. As this variable is dependent only on the therapist, it is not calculated based on the QuestionnaireResponse resources.

Interpretation of the answers to a mental health screening questionnaire can not be performed in a generic fashion towards all questionnaires. Even though the questionnaires follow the patterns outlined in section 3.3, the methods used for interpreting the results vary between questionnaires. As an example we can consider the MADRS and PHQ-9 questionnaires. An answer sum of 20 in MADRS entails ”possible or mild depression” [61]. In the PHQ-9 questionnaire, a sum of 20 is considered severe depression [21]. This is a simple example, and there exist other methods for interpretation. Another example is interpreting the answer differently depending on the which question was answered, which is done in the MADRS questionnaire for the suicidality question.

As a solution to not having a single method for interpreting QuestionnaireResponse resources, we provide a fast way to tailor the application towards any iCBT programme. By using the interfaces shown below, functions can be created to interpret the results of any standardized questionnaire. All necessary information for these functions is available in the list of QuestionnaireResponse objects. Default implementations are included in the artifact, based on the MADRS questionnaire. Replacing the functions will require programming skills, but will not require understanding of the artifact beyond the interfaces.

#### 4.5.1 Replaceable functions

The artifact supports any functions implementing the interfaces below, to calculate the status variables. These functions should be tailored to the iCBT at hand, with the cooperation of therapists in order to provide optimal results. The following list provides a description of the status variables that needs to be tuned to the treatment:

- The warning variable should only trigger for patients who require urgent attention, and provide some short description on why. As these functions actively suggests patients to prioritize, this is possibly the most important status variable. The implementation of the function calculating these variables should be understandable for the therapists. If the therapist knows the limitations of the warnings, they are better suited to make decisions on how to apply the presented information.
- The progression variable should provide some overall measure of how the patient is progressing, based on the data in the QuestionnaireResponse resources. In order to provide a chart of the overall state of the patient population, the possible strings for the progression variable are limited to ”improving”, ”steady” and ”declining”.
- The flag variable acts as a more generic variable which can be tailored to the iCBT situation. As the concept of flags is very versatile in this setting, there are multiple options for how to tailor these to a specific

treatment. Some examples of potentially useful functions for flag variables are provided at the end of this section.

- The urgency score variable provides a way for therapists to automatically rank patients based on their need for care, similarly to triage. This function is probably the hardest to implement correctly in practice. Correctly estimating each patient's urgency might also require more data than just information from questionnaires. The default function implemented in the artifact displays the latest MADRS sum. Assuming the urgency score can be calculated with a satisfactory degree of accuracy, this variable can enhance therapist workflow. The possibility of sorting the table in the Master View on this variable provides the therapist with an option to assist the most urgent patients first. The goal of this variable is presenting the urgency information in advance, compared having it after processing each patient.

The "Time spent in program" variable is calculated as max time elapsed since any QuestionnaireResponse resource was authored, for each patient. Replacing the default functions will require changing one line of code, excluding the code required for the new function. As the status variables are calculated on the application server, the functions are currently limited to C#. The functions for calculating the variables are represented as methods on C# objects. The interfaces for these objects are displayed in figure 4.5.

```
public enum ProgressionRepresentation{ improving, steady,
                                     declining, blank, error };

public interface IAggregationFunction{
    ProgressionRepresentation aggregate(List<QuestionnaireResponse> QRs);
}

public interface IFlagFunction{
    List<string> calculateFlag(List<QuestionnaireResponse> QRs);
}

public interface IWarningFunction{
    List<string> calculateWarning(List<QuestionnaireResponse> QRs);
}

public interface IUrgencyScoreFunction{
    int calculateUrgency(List<QuestionnaireResponse> QRs);
}
```

**Figure 4.5:** The interfaces provided for calculating status variables

The input of the functions is the full list of QuestionnaireResponse resources for a patient. The "blank" field of the ProgressionRepresentation enum is displayed in the GUI as the empty string. This is intended for patients with insufficient data. The "error" field can be used for development purposes. The

possible strings for the progression variables are restricted to the fields in the enum to enable color coding of the pie chart in the Master View.

The flag function provides a way to incorporate any relevant information in the patient table. Examples of potential use cases:

- Predictive analytics: As the list of QuestionnaireResponse resources includes all the previously recorded answers to questionnaires for the patient, a pattern-recognition algorithm could be plugged in here to flag potentially interesting patients. Predictive analytics has previously not been used much in clinical dashboards [38], and this could be interesting to further investigate. This concept could also be applied to other features of the dashboard. Combining predictive- and visual analytics could for example enable therapists to see forecasts for patient progression.
- Sudden changes: Large delta values between the latest two data points could warrant therapist attention. The current default flag function does this for the MADRS questionnaire. Specifically, the category of the largest delta subject to  $\text{delta} \geq 2$  is returned. This threshold is probably too low for use in practice, but demonstrates how flags can be presented in the artifact.

## 4.6 Scalability

As one of the objectives of the research was to assist therapists in treating more patients, discussing scalability is relevant. We can consider scalability in terms of number of therapists, number of patients per therapist, and data points per patient. The design considerations has been described in previous sections. The following sections will discuss the technical scalability of the artifact.

As the resources used in the artifact were generated, we were able to simulate scaling by generating more data. The final version is designed to support hundreds of patients per therapist. The exact amount will depend on server hardware. The artifact functions without any user experienced latency due to data processing, at the scale of  $\approx 150$  patients with  $\approx 200$  average resources per patient, running on a laptop. This is discussed further in section 5.2.

Throughout the development process, the resource server contained approximately the same amount of resources. There were no noticeable performance problems initially. As more features were added, latency problems were noticed. The main performance bottleneck was identified to be interaction with the FHIR server, where retrieving the resources for one patient could take multiple seconds. The introduction of the status variables (described in section 4.5) lead to considerable user-experienced latency, as a large amount of resources were required to calculate these variables. As the latency from loading the Master View exceeded a reasonable threshold (2-3 seconds), some steps were taken to pre-calculate, cache, and optimize the calculations. The goal was for the processing time to not affect user experience.

#### 4.6.1 Scaling measures to improve user-experienced latency

Initially, the application server retrieved all relevant resources from the FHIR server on every request. This approach proved infeasible when several seconds were required to load the main page. Caching the resources in the database of the web server was the first optimization measure. This provided a slight performance increase averaged over multiple users, but user experience was still impacted significantly. As common use cases for the artifact entail a low amount of users, the cached entries would be out of date often enough for users to notice performance problems.

During this optimization process, we identified HAPI JPA paging as one of the performance problems. The default configurations for the resource server is to return pages of max 10 resources on a search. The effects of this is described in section 5.2. We disabled this feature for faster development. As other FHIR servers might use paging, it is still supported in the artifact.

For the second optimization measure, the logic for processing the bulk of the resources were separated from the logic called on requests. This process could then be scheduled to run at fixed intervals, resulting in pre-calculated data being available for the views. This approach led to the user not noticing any latency when using the artifact.

The optimization measures for the artifact solved the problem of user experienced latency. Other measures are available for further optimization. These were not explored further in the artifact, as the implemented measures achieved satisfactory performance. Even though these options were not implemented, they can be discussed based on the findings in the design process.

As the user experience is fluid with the optimization measures mentioned above, further measures considers optimizing the update process. As described in section 5.2 loading all the required FHIR resources for updating the data takes a significant amount of time. An empirical test, described in section 4.6.1, measured a run time of  $\approx 10$  minutes for executing the update process once. Given similar runtime on a server, a possible interval for updates could be every hour.

#### 4.6.2 Possible scaling measures to improve update time

The following optimization measures have not been implemented in the prototype, but present natural steps to take for further optimization of the dashboard application. The measures taken in in section 4.6.1 enables the dashboard to scale well enough for use at the current iCBT scale. Further optimization is possible, but given limited development time, we did not implement them in the artifact.

FHIR Subscription [77] offers significant potential for optimization. Storing the data in an efficient format in the database would also help this approach. FHIR Subscription would enable the artifact to receive new resources when they are posted to the resource server. The resources could such be stored in

the database of the application server, and processed at regular intervals. This approach would eliminate the impact of network latency, due to the callback nature of the Subscription. An argument could also be made for processing each resource every time the FHIR Subscription returned a resource. This could be possible, but would lead to increased computations if multiple QuestionnaireResponse resources were received for the patient over the update interval. Using FHIR Subscription would also reduce the load on the resource server, which could be beneficial if this server interacts with multiple applications.

An optional approach to using FHIR Subscription could be to request resources based on date/time. This method implies querying the database of the web server for the date of each resource type, every update. The query would need to find the latest date for each resource type, for every patient, at every update. This method would provide a similar performance improvement compared to FHIR Subscription, but is potentially harder to implement. Finding the latest date for each resource type would either require a substantial amount of extra database operations, or organization of the database structure to efficiently find these dates. For the same comparable efficiency to FHIR Subscription, the dates would need to be retrievable in constant time.

Requesting new resources based on date was initially implemented for the artifact, but without optimizing storage format and database structure only minor improvements were observed. With a naive database structure, getting the latest date for each resource type for each patient requires checking all resources in the database. FHIR Subscription provides the performance benefits regardless of database structure, which indicates it might be a better solution.

The previously mentioned measures to scale the application with respect to update time were focused on reducing the amount of resources sent over the network. Another possible measure for achieving this could be to make the requests from the application server asynchronous. This could also be combined with the date filter method. Using asynchronous methods to make the requests would mitigate the effects of network latency, but would also add complexity to the code. Asynchronous requests would not solve the issues related to paging, as the links to the next page is returned with responses.

Other alternatives for optimizing performance could be to modify the resource server. As the artifact was made to be compatible with existing FHIR servers this was not explored beyond disabling paging on the HAPI JPA server. For applications requiring especially large volumes of data, however, this might be necessary.

Extra considerations could be required if the resource server is shared between many heterogeneous client applications. In this situation, there is likely an abundance of resources which are irrelevant to the artifact. Making more specific queries could be a possible remedy for this problem. The filtering of resources could then be done on the resource server, rather than the web server. Another possible solution for dealing with too much response data could be to use GraphQL [78]. GraphQL provides methods for dealing with large objects



containing little useful information. Implementing this would, however, require substantial changes of the source code, and was not done in the artifact.

## 4.7 Security considerations

As the objectives and research questions of the design science research does not consider security, this has not been implemented in the artifact. For the artifact to be used in a real setting, security would need to be implemented first, as medical data is sensitive. Some thoughts about how this can be done are presented here.

In a practical setting, a security protocol should be implemented on the resource server. This could further be layered with security on the web server of the dashboard. As therapists in practice might not have access to the same patients, the therapists security principals have to be mapped to their patients. This could in theory be performed either on the resource server or the web server, or both. As the artifact processes all the patients in one batch, mapping the therapists to their patients is probably easiest achieved on the web server.

In order not to interrupt the workflow of the therapists, it would be beneficial for the application to have external authentication, for example OAuth. If therapists can log in with the same credentials as on their other systems, this would make the system easier to access. Depending on the environment for deployment, this authentication could possibly be automatic when the therapist logs on to the computer.



## 5 Demonstration and Evaluation

Evaluation of the artifact is an essential part of design science research. Peffers et. al. prescribes demonstration and evaluation of the artifact as the fourth and fifth activities of design science research [14]. These activities are related, as the evaluation of the artifact depends on the results observed from the demonstration. During evaluation, the observations from the demonstration should be compared to the objectives for a solution.

The demonstration activity was done by presenting a running version of the artifact for two experts in psychology and psychiatry. The artifact was then evaluated through a semi-structured interview with these experts. Following this, the technical scalability of the artifact is evaluated through empirical measurements of run time. Lastly, we evaluate how the artifact compares to dashboard design guidelines from a meta study on clinical dashboards.

### 5.1 User evaluation through tests and semi-structured interviews

As the objectives of the artifact take a qualitative nature, semi-structured interviews were used to gather qualitative data about the design. While qualitative research methods originate in sociology and anthropology, they have also been applied in the software engineering domain to gather qualitative data [79]. Using interviews is beneficial for collecting data which cannot be obtained using quantitative methods [79].

We performed the interviews with two experts in psychology and psychiatry. Semi-structured interviews were chosen because we only expected to have the opportunity to interview each expert once, and because we only had two interviewees. This interview format enables the interviewees to provide any relevant information, and not just answers to the questions. One of the experts has practical experience from working in iCBT. Ideally we would have conducted the interviews with more experts, but given the time frame and scope of the project, we were able to recruit two qualified experts. Figure 5.1 shows the interview setup. The interviews were structured as following:

1. A short presentation of the artifact. This consisted of a brief explanation of the artifacts features.
2. Next, the interviewee would test the artifact while presenting his/her opinions underway.
3. After testing the application, the interviewees were asked the questions presented in Appendix A. These questions take an open ended form, and they are meant to provide more information than general usability. As the interviewees were experts, the questions were formed to gather the necessary information to evaluate the artifact and gain qualitative data towards answering the research questions.

The first two parts of the interviews entail the demonstration activity of design science.



**Figure 5.1:** Interview setup

The qualitative data obtained in the interviews were used both to evaluate the design science artifact, and to gain insight into further work for data analytics in iCBT. The interviews helped determine which features of the artifact were useful for guided iCBT, as well as which additional features could be useful. This provided us with qualitative data which can be used to answer the research questions. Most of the additional features identified in the interviews fall outside of the scope of the artifact presented in this thesis, and is mentioned under further work in section 7.2.

### **5.1.1 Interview results**

The interviews were recorded on video, and then transcribed. Interviewee 1 is a psychiatrist with experience providing guided iCBT. Interviewee 2 is a researcher and full professor with a background in psychology. As interviewee 2 has limited experience with iCBT, questions 12 and 16 were skipped during the second interview.

During the first two phases of the interview, we observed some confusion towards certain parts of the user interface. Interviewee 1 was initially confused by what happened when various things were clicked. In the Detail View

specifically, how to open/select charts was a point of confusion. Interviewee 2 expressed confusion about which patients were considered by the pie chart. Both of the interviewees understood how the application worked after testing it further. A summary of the answers to all questions in the interviews are presented in appendix A.

Overall, the experts expressed positive opinions about the presented functionality. Multiple new features were identified as requirements for use in practice. As these features fall outside of the scope of the project, they are further discussed in section 7.2. Many features of the application was pointed out as useful. Furthermore, potential benefits of the displayed functionality was highlighted. These are discussed further in section 6.

## 5.2 Scalability measurements for evaluating system performance

This section will present measurements of system performance. Through the interpretation of these measurements, the main performance bottlenecks will be identified, and the consequences of these discussed. All measurements presented in this section consider update time as described in section 4.6. The therapists using the dashboard will not experience any latency, due to the measures described in section 4.6.1.

In order to gain a detailed view of how the application scales with respect to the amount of data points, C# methods in the artifact were timed during the update. The empirical measurements show that interaction with the resource server is responsible for the majority of latency. The functions were logged through the use of the C# Stopwatch class [80]. The measurement accuracy can be affected slightly by other processes running on the computer. The experiment was run twice with the same resources, to observe the effect of resource server paging.

For the experiment, the HAPI JPA reference implementation cloned in November 2018 was used. The Apache Tomcat [81] application server, version 7.0.90 was used for running the FHIR server. The specifications of the laptop used to run the experiment were the following:

- **Operating System:** Microsoft Windows 10 Home
- **Processor:** Intel Core i7-6700HQ CPU @ 2.60GHz, 2601 Mhz, 4 Core, 8 Logical Processor
- **RAM:** 16 GB

When performing the measurements for the single patients, the code for retrieving the FHIR resources (Observation, QuestionnaireResponse, Questionnaire) only includes the time for making requests to the resource server, and for assembling the list of resources. For the measurements of higher abstraction level methods, displayed in tables 5.3 and 5.5, other server code has an effect on

the run time. Examples of such code would be running the status variable calculations, format wrangling, and database operations. The measurement for one patient’s QuestionnaireResponse resources were performed with the following code:

```
stopwatch.Start();
List<QuestionnaireResponse> QRs = new List<QuestionnaireResponse>();
Bundle results = client.Search<QuestionnaireResponse>(new string[] {
    "subject=Patient/" + id });

while (results != null){
    foreach (var entry in results.Entry){
        QuestionnaireResponse QR =
            (QuestionnaireResponse)entry.Resource;
        QRs.Add(QR);
    }
    results = client.Continue(results);
}
stopwatch.Stop();
```

**Figure 5.2:** Code snippet for requesting all QuestionnaireResponse resources for one patient

Similar code is used to measure the run time for retrieving the other resource types for each patient. The code snippet in figure 5.2 retrieves and unpacks the QuestionnaireResponse resources for one patient and loops through all pages if paging is enabled. The elapsed time between "stopwatch.Start()" and "stopwatch.Stop()" corresponds to the QuestionnaireResponse rows in tables 5.2 and 5.4. As such, the latency induced by REST and the deserialization with the C# FHIR client [69] is included in the measurement. Using the while loop to handle paging is the prescribed method in the official documentation for the .Net FHIR library [82].

For the experiment, we used 147 generated patients with their corresponding data, as described in 4.4. Table 5.1 shows the amount of resources retrieved for one update. All resources were retrieved once, except the Questionnaire resources. The implementation of the prototype requires the QuestionnaireResponse resources before requesting the Questionnaire resources. Consequently, additional logic is required in order to not request each Questionnaire more than once. This would provide a minor improvement to update time, but is not implemented in the final artifact. The other resources in table 5.1 were requested once per resource.

Total	QuestionnaireResponse	Observations	Questionnaire	Patient
22023	7705	14024	294	147

**Table 5.1:** Total resources retrieved in one update

All functions measured here runs on a scheduled server process, as described in sections 4.1 and 4.6. For the user, all interactions with the server will happen close to instantly. When performing the measurements, the application server and the resource server were running on the same computer. This might not be the case in a real situation, which could lead to a longer run time due to network latency. The following tables show the measured run time of the functions considering FHIR resources:

	$\mu$ time	$\sigma$ time	$\mu$ resource count	$\sigma$ resource count
Observations	0.90s	2.48s	95.40	163.29
QuestionnaireResponses	0.34s	0.10s	52.41	23.92
Questionnaires	0.41s	0.04s	2.00	0.00
Total patient	3.05s	3.21s	149.81	166.71

**Table 5.2:** Single patient measurements

	process runtime
Get all patients	4.79s
Update all QRs	86.64s
Calculate functions	0.05s
Complete update	541.10s

**Table 5.3:** Measurements across all patients

The effect of paging can be observed through all operations interacting with the resource server except for Questionnaires. This resource type is unaffected as only two questionnaire resources are requested per patient, and this is below the page cap. These questionnaires are also requested in sequence, which explains why Questionnaires takes more time than QuestionnaireResponses in table 5.2. By enabling paging on the resource server, the run time of the update process increased by  $\approx 347$  seconds, equivalent to 56%. From the runtime measurements, we can also observe that the time required to calculate the models for the Master View takes  $\approx 50$  milliseconds.

	$\mu$ time	$\sigma$ time	$\mu$ resource count	$\sigma$ resource count
Observations	2.42s	4.01s	95.40	163.29
QuestionnaireResponses	1.22s	0.52s	52.41	23.92
Questionnaires	0.41s	0.04s	2.00	0.00
Total patient	4.52s	4.69s	149.81	166.71

**Table 5.4:** Single patient measurements, with paging

	process runtime
Get all patients	6.34s
Update all QRs	216.05s
Calculate functions	0.04s
Complete update	888.39s

**Table 5.5:** Measurements across all patients, with paging

As the Master View of the dashboard requires all `QuestionnaireResponse` resources from all the patients, calculating all the status variables will take multiple seconds even with a small amount of patients. As displayed in tables 5.3 and 5.5, the calculations does not take a significant amount of time, but the overall process takes multiple minutes in this experiment. Requesting all the data from the resource server on each request to the application server would require a long waiting time for the user, and would also not scale well with the amount of therapists. This flow could for example be the case when performing all calculations on the client with JavaScript. While this approach could be acceptable for other types of applications, we have shown that this is inefficient for dashboard applications which requires a large amount of data.

### 5.3 Heuristic evaluation with dashboard design guidelines

At the initial stages of designing the artifact, we reviewed the literature to identify appropriate methods for presenting the data. Section 2.5 presented the specific systems identified during this review. Meta studies on clinical dashboards were also identified, among them Khairat et. al. [44]. This study presents 10 guidelines to aid development, testing and implementation of clinical dashboards. While the ICU setting is mentioned specifically in the background section of this article, the literature review considers any healthcare setting. The guidelines presented by Khairat et. al. were used to guide the design of the artifact. In this section we will discuss how the design science artifact conforms to these guidelines.

**G1: Apply realistic techniques to enhance mapping of data elements to visual objects.**

The interactive spider chart is the only novel visualization in the artifact. As the other charts can be considered common visualizations, we regard these techniques as realistic. During the user tests described in section 5.1, none of the interviewees indicated confusion towards the charts in the Detail View. There were some initial confusion about which patients were shown in the charts of the Master View. Both interviewees indicated that the interactive spider chart was useful, and that a spider chart is appropriate for visualizing the screening questionnaires.

**G2: Minimize user actions to accomplish a goal.**



This guideline was central for the artifacts design. As the artifact is designed to conform to Shneiderman’s mantra [48], the process of reaching any view state will consist of tunnelling down, or going back up. The decision to persist the selection of charts for each patient further contributes to this guideline. As a therapist will not need to re-select all the relevant charts each time a patient is clicked, this reduces the amount of required user actions.

**G3: Provide flexibility in the ways to achieve the same goal.**

As the artifact is designed to facilitate analytics, the users can choose between multiple options for what to do at any time. We can provide two examples to illustrate this:

- Goal: Find out if a patient has improved or declined on the "reduced sleep" category of MADRS. This can be achieved in two ways. After opening the interactive spider chart for the patient in question, the therapist can slide the mouse over the line chart to observe change in the spider chart. Alternatively, the axis for reduced sleep can be clicked to see the data points in this category as a line chart.
- Goal: Find the group of patients with declining total MADRS score. The table in the Master View can be sorted to display this group of patients on the top. This can be achieved either through clicking the red slice of the pie chart, or through clicking the "Progression" column in the table.

**G4: Provide functionality to represent additional information.**

The therapist has a high degree of control over which information is shown on the dashboard. A therapist can choose which observations are displayed in the Detail View. Similarly, she has the option to select relevant axes of the spider chart as described in guideline three.

**G5: Spatially organize the visual layout.**

The bootstrap grid system [83] is applied to the GUI in the Detail View in order to ensure vertical alignment of the visual elements. This view is first split vertically in in halves. The right side of the view is reserved for line charts. The left side is further split vertically, below the name of the patient. Overall, this divides the view into three columns of visual elements, where the elements in one column represents one functionality.

The Master View is organized with the most important information on the top of the view. The number of patients could potentially grow to extend the view multiple times. Displaying the visualizations on top, and providing methods for placing relevant patients on top of the table enables the therapist to maintain an overview of a larger patient population.

**G6: Consistently apply design choices.**

Color coding for the visual elements is applied carefully throughout the artifact. The color red is only used to indicate problematic patients in the pie chart. Similarly, green is only used to indicate something going well. The default neutral color is blue (steel blue) for both views. The buttons used in the artifact also have the same theme and similar coloring.

**G7: Place minimal cognitive load on the user.**

The cognitive load will depend on the goal of the therapist. If a therapist wants to compare patients at a detailed level, the application currently does not support this well. When comparing data for a single patient, however, all necessary information is available through one view. The artifact also presents a high level overview of the group of patients in the Master View. Compared to the current systems in eMeistring, as described in question 4 of Appendix A, this is a significant improvement.

**G8: Provide users with information on alternatives when several actions are available.**

The artifact does not fulfill this guideline. Ideally, with more development time, implementing info buttons could have helped users understand the options better.

**G9: Remove extraneous or distracting information.**

Everything displayed in the GUI of the artifact fulfills a function. The users have a high degree of control over which data is displayed. In the Detail View, therapists can control which data types they are interested in, with the exception of the background information. For the Master View, enabling the users to select the data in a similar manner could be beneficial, as presented in the second interview.

**G10: Consider means to reduce the data set.**

The artifact does not display any means. This could potentially have been beneficial for the numerical variables displayed in the master view (Urgency score, Time spent in program). Other techniques are applied to reduce the data set for other variables. The pie chart reduces the categorical progression variables for all patients to three percentage values. Similarly, sums are used to reduce the questionnaire data in the Detail View, with the details on demand in the interactive spider charts.

## 6 Findings and Discussion

Hevner et. al [15] describes guidelines for design science research. The presented framework specifies how design science should both solve problems in the artifact domain, as well as provide research contributions in the form of additions to the knowledge base. In this section we will recap the contributions produced by the design science research process, both towards the knowledge base and the iCBT domain.

### 6.1 Contributions to the knowledge base and answers to the research questions

The contributions to the knowledge base considers the new information discovered throughout the design process. To a large degree, this information overlaps with the answers to the research questions. Presented below is the answers to the research questions, as well as a discussion about which new knowledge was discovered.

#### **RQ1: What data representation of mental health data in iCBT can facilitate analytics?**

During the semi-structured interviews, both interviewees expressed that the data in the artifact was useful, although they would require more parameters for practical use. Some of the discussed data types were qualitative in nature, but multiple quantitative measurements were also mentioned. Of the quantitative measurements, both experts mentioned questionnaires and physiological measurements. The artifact presents ways to store, process and visualize both of these data types.

Other data types were also mentioned, such as life events and messages. As these data types can not be represented as numerical or ordinal data they fall outside of the scope for the artifact, but would be useful to consider for further research. The modules/exercises were also discussed. Unpacking the currently text-based modules into structured data types for presentation is a large but useful task, which can be researched further.

Considering the psychometric evaluation screenings, we identified the patterns described in section 3.3. Storing of these questionnaires in FHIR format is beneficial for system interoperability, as well as for further processing. Examples of how to generate and process these resources have been presented. Furthermore, we represent the psychometric screening questionnaires over time as unevenly spaced multidimensional time series. Through this representation, methods for visualizing multidimensional data can be applied to present these series.

For visualization of physiological measurements, a similar approach as to the questionnaires can be applied. Unless the data has high velocity, each data point can be represented by the Observation resource. The Observations can further be converted to single dimensional time series. If the physiological measurements

are single dimensional (e.g. not blood pressure), they can be presented in the same way as aggregates over questionnaires.

**RQ2: Using the data representation from RQ1, how can the data be visualized for therapists to gain actionable insights into the state of a single patient, or a population of patients?**

The answer to this research question can be divided into an answer for the patient group, and an answer for the single patient. The artifact presents examples for visualizing data for both the single patient and a group of patients.

Considering the single patient, we have provided examples of how visualizations can be applied to the data from RQ1. For any series of numerical observations available for a patient, regular line charts are appropriate visualizations. The option for the therapist to select which measurements should be displayed for the single patient was determined to be useful by the experts we interviewed. For the screening questionnaire data, both experts indicated that the line charts with the sum over the answers were useful presentations. Both experts also indicated that these charts could be used by a therapist to make decisions.

Additionally for the single patient, we developed an interactive spider chart to visualize the patient filled screening questionnaires in the form of multidimensional time series. The interactive spider chart was determined to be a useful representation of the questionnaire data by both interviewees. This visualization can provide the therapist with fast overview, and can give a full presentation of the data within the series of screening data. To our knowledge, this interactive visualization method has previously not been applied for visualizing multidimensional time series.

For the patient population, we developed a pie chart dividing patients into three categories depending on their symptom progression, measured by the screening questionnaires. Both interviewees indicated that the pie chart was useful. Interviewee 2 indicated this visualization could be useful for comparing sub-populations.

**RQ3: How can decision support functionality be implemented to assist therapists in understanding each patients required urgency of treatment in a population of patients, based on the data from RQ1?**

In order to provide an overview of the patient population, we implemented a set of status variables for each patient. The variables presented in the artifact are: "Time spent in program", "Flags", "Progression", "Warnings", "Last Checked" and "Urgency Score". These were described in detail in section 4.5. The variables can be displayed in a table to gain a fast overview of a small to medium patient group. By implementing functionality for filtering, sorting and searching the table, overview can be gained for a larger patient group. By further applying interactive visualizations to these variables, functionality for visual analytics can be provided.

The status variables are calculated based on self-reported values from the screening questionnaires for each patient. Both interviewees indicated that the status variables were useful, and suggested additional variables to be displayed for each patient. The status variables follow simple interfaces outlined in section 4.5.1, and are easily adaptable to different iCBT treatments.

**RQ4: How will a dashboard implementing the techniques from RQ2 and RQ3 scale when using the HL7 FHIR format for medical interoperability?**

The steps taken to scale the application to more patients per therapist with respect to the GUI has been described above. The findings with respect to performance are presented here. As dashboard applications considers a multitude of data points per patient, considering scalability is relevant in order for performance issues not to negatively impact user experience.

Scalability considers how the system’s performance scales with respect to the amount of therapists, patients, and resources. We found the HAPI JPA reference implementation of the HL7 FHIR DSTU3 standard not to scale well with regard to this, as described in section 5.2. While this particular implementation does not scale well, this is not necessarily an indication of the HL7 FHIR standard not scaling. We present several possible measures to consider for dealing with the scalability problems in section 4.6. The HL7 FHIR standard provides many positive properties, such as system interoperability and a common interface towards the data. As the performance problems can be dealt with, the benefits of using the FHIR standard outweighs the extra efforts required for optimizing performance.

## 6.2 Contributions to the domain

The artifact provides solutions to relevant problems in iCBT. We designed the design science artifact with the goals of solving current problems discovered in practical iCBT. Through evaluation by experts in psychology and psychiatry, the artifact was determined to provide solutions which would be beneficial for guided iCBT in practice. Within the scope of this project, as outlined in section 1.3.3, the artifact displays satisfactory results.

Expert evaluation identified the functionality of the artifact as useful for both current iCBT as well as potentially regular CBT. The development of the artifact was motivated by real clinical needs within guided iCBT. Through the development and evaluation of the artifact, the following problems and solutions were identified.

- Status variables for a group of patients were identified to be useful for gaining an overview of a group of patients in iCBT.
- Visualizing quantitative measurements for patients can enable therapists to faster gain an overview of their patients.
- Additional tools for comparing patient sub-groups can be beneficial.

While not entirely unknown previously, it is also worth mentioning the importance of further structuring the data in iCBT to enable analytics. Specifically important is the content of the treatment modules and patient filled exercises. There is a lot of useful information in these modules, but currently free text is the main way to represent the answers to the exercises.

## 6.3 Validity

Internal and external validity of research is most clearly defined in studies applying statistics. Lukyanenko et. al. [84] addresses validity of IS design science research through instantiation validity. Instantiation validity "refer[s] to the validity of IT artifacts as instantiations of theoretical constructs". The article discusses the lack of guidelines in available in the literature for demonstrating the validity of design science research. Several threats to the instantiation validity inherent to the design of software systems are presented. Examples of these threats are:

- **Instantiation space:** The set of possible variations in the features of software.
- **Artifact complexity:** Software often requires additional properties in order to test one feature. The article presents the example of user interface design often requiring realistic data for the user to interact with.
- **Artifact cost:** Designing a software artifact is an expensive process. A consequence of this is that often researchers can only develop one artifact, resulting in comparison between multiple potential solutions not being an option.

Lukyanenko et. al presents additional threats, and concludes that further research is required on developing guidelines for IS research to mitigate these threats. Hevner et. al. [15] discusses validity in terms of research rigor. The article presents the following view: "[...] rigor is derived from the effective use of the knowledge base-theoretical foundations and research methodologies. Success is predicated on the researcher's skilled selection of appropriate techniques to develop or construct a theory or artifact". Based on the articles [84][15] presented in this section, giving a precise answer about the validity of the design science research is not feasible. Below, we will argue how the techniques applied during the research process enhanced the validity of the design science research project. Because of the nature of the design science research, we will apply the terms internal and external validity in a more general sense.

### 6.3.1 Internal validity

The user evaluation through the tests and semi-structured interviews identified that the research presents acceptable solutions to problems in one case of iCBT. During these interviews, the artifact was mainly compared to the eMeistring treatment programme. The techniques presented by the artifact was well received by the experts, as presented in section 5.1 and Appendix A. Testing the artifact in a practical setting would be beneficial in order to give a more definite answer on the applicability of the artifact to eMeistring.

### 6.3.2 External validity

In this section we will describe the artifacts ability to extend beyond the case of eMeistring. The goal of the design science research was to solve relevant problems in iCBT. As such, we will consider the artifacts applicability towards other iCBT programmes.

The dashboard design guidelines presented in section 5.3 considers any medical dashboard application. The artifact mainly conforms to these guidelines, which indicates the artifact could be extended to situations beyond the eMeistring case. The use of the HL7 FHIR standard ensures the data considered by the artifact has a general nature. The design science research of this thesis considers visualization of questionnaires, and numerical observations. As these data types are relatively generic, the artifact is likely adaptable to other iCBT situations. The constraints on eligible questionnaires, outlined in section 3.3, are fulfilled for most standardized questionnaires used for iCBT. All the questionnaires mentioned in table 2.1 with psychometric properties follow these constraints. We have not been able to identify any standardized mental health screening questionnaires not conforming to the first two constraints in section 3.3, and the third constraint is optional for use in the artifact, although it is recommended.

Multiple additional required features were identified during the evaluation of the artifact. As different iCBT treatment programmes have differing nature, one generic approach can not be expected to solve all problems in this domain. Considering the features implemented in this research project, we argue that these features has the potential to be extended beyond the case of eMeistring.

## 6.4 Reflection over tools and frameworks

The tools used in the artifact changed throughout the development of the design artifact. A description of the tools and why we chose them is given in sections 4.1 and 4.2. Most of the tools worked reasonably well for their intended purpose in the artifact. This section will present some of the issues identified with the tools, as well as general thoughts about each tool. The tools are grouped somewhat more compactly than in sections 4.1 and 4.2, as some of the findings are shared across multiple tools.

- **Matplotlib, Seaborn and Dash:** These tools provided the expected functionality. As described in section 4.1, the tools allowed us to get a rapid first prototype. This was useful in order to further plan the design and development of the artifact. We, however, do not recommend using these tools for a fully developed production web-app.
- **SMART on FHIR JavaScript and Python clients:** Both of these libraries are popular FHIR client implementations for their respective languages. As SMART on FHIR focuses on providing security, and we did not use this functionality in the artifact, we can only present insight into some aspects of the libraries. Our main impressions of the libraries considers the lack of documentation. Some simple tutorials are available on the SMART on FHIR website [56], but these only provide basic examples. Apart from this, we could only find auto generated documentation

with limited usefulness. As such, we found it necessary to read the source code of the libraries in order to perform certain operations. The lack of documentation can be attributed to the tools being fairly new.

- **ASP .NET MVC:** This heavyweight web framework provided a lot of options. We were only able to utilize a small fraction of the potential functionality available in the .NET frameworks. The general MVC structure, and ease of extensibility was useful when designing the artifact. In hindsight, maybe a lighter framework would have been better suited for the artifact, but ASP MVC provides a good foundation for a larger project.
- **MongoDB:** Due to the heavy use of FHIR resources in the artifact, MongoDB worked quite well. The document based database worked well both for the initial caching of FHIR resources, and the eventual storing of the models. Being able to store models in a format which can be serialized to JavaScript objects was also useful. We noticed one problem with this approach, when using the MongoDB C# driver [73] in conjunction with the C# FHIR client [69]. When retrieving resources from the FHIR server, the C# FHIR client would deserialize these to C# objects. This is necessary in order to manipulate these objects on the applications server. However, when storing these objects in MongoDB, the BSON would include nulls and empty lists for all the empty fields in the resource. The MongoDB C# driver does not have any good methods for fixing this problem globally, as far as we could find. The issue is described in detail in the official MongoDB Jira [85]. In practice, this problem had no effect on the logic for storing and retrieving the resources in the artifact, and only impacted the required size for storing the BSON.
- **HAPI JPA:** The scalability problems of the HAPI JPA Java reference implementation of FHIR is discussed thoroughly in sections 4.6 and 5.2. For use cases requiring a small volume of transmitted resources, the HAPI JPA implementation is an acceptable solution. Setting up the resource server was fast, and relatively unproblematic. For larger scale projects, we recommend modifying the server or using another implementation.
- **Synthea:** The Synthea Patient Generator provided a good starting point for patient resources. The generated resources has a general nature, and did not fulfill all our requirements for resources. This is mainly due to our use case being very specific. For getting started with FHIR, the generated resources provided a good baseline.
- **Highcharts and D3:** Both of the JavaScript based visualization libraries provided high quality solutions for the prototype. Highcharts provided out of the box solutions, and D3 a high degree of customizability. Both libraries are well documented and commonly used. No significant problems were encountered when using these libraries, and we recommend using them for any data visualization project. If the non-commercial clause of Highcharts is a problem, the Plotly.js library [86] provides similar solutions.



Some of the tools worked better than others, but the overall combination of the tools presented an acceptable result. The complete tech stack is heterogeneous with respect to programming languages. As the FHIR protocol is based on REST, we did not experience any major problems from this. Because of the REST protocol, many of the logical modules presented in the architecture section 4.1 are replaceable. Many of the libraries could have been replaced, with the resulting artifact still providing the same functionality. Even though we identified problems with some of the tools, it is unclear which other tools would have provided better results.

## **6.5 Communication**

The first five activities of design science research has been presented in the previous chapters. The communication activity as presented by Peffer et. al. should "communicate the problem and its importance, the artifact, its utility and novelty, the rigor of its design, and its effectiveness to researchers and other relevant audiences, such as practicing professionals, when appropriate". This thesis presents all of these aspects, primarily towards researchers. During the evaluation of the artifact, one relevant practicing professional within iCBT also received a thorough presentation of the artifact. We also plan on publishing a research paper on the results in this project, in a scientific journal.



## 7 Conclusion and Further Work

This thesis presents a dashboard application designed to provide therapists in guided iCBT with decision support and visual analytics. The dashboard- and visualization functionalities presented in this thesis can help guide further research on visual analytics in iCBT. This section will summarize the design science process and results, and present suggestions for further research on presenting patient data in iCBT.

### 7.1 Conclusion

Through the design science methodology, we designed, implemented, and evaluated an artifact created to solve relevant problems regarding a therapist's overview of patient data in iCBT. Specifically, the artifact considers standardized mental health screening questionnaires and numerical time series. We have presented methods for providing therapists with overviews towards single patients, as well as for a group of patients.

As described by Peffers et. al. [14], the development of an artifact should be a search process using existing theory to develop a solution to a defined problem. During the process of designing the dashboard application, multiple findings were uncovered. Both simple (line-, bar- and pie charts) and complex visualizations (interactive spider chart) were used to display quantitative data for patients in iCBT. A novel way to visualize mental health screening questionnaires over time was developed, in the form of an interactive spider chart. Furthermore, scalability problems were identified with the HAPI JPA server reference implementation of the HL7 FHIR standard. Measures to improve user-experienced latency have been implemented, and further scalability measures have been discussed.

### 7.2 Further Work

In order for the application to be used in a practical setting, multiple features would need to be implemented. The scope for the design science artifact, described in section 1.3.3, does not include security or unstructured data from the treatment modules. Both of these aspects should be implemented before testing the application in a practical setting. Furthermore, the extra features suggested by the experts during the evaluation would be beneficial to implement. These include:

- Patient reported events.
- Messages from the patients.
- Tools for comparing patients.
- Tools for analyzing sub-groups of the patients.
- Options for the therapists to manipulate the line charts in the following manners:
  - Starting y-axis at 0.
  - Align charts to a common time axis.

- Zoom in to a subset of the data points.
- Combine two charts to display ratio over time.

The interviewees also asked for additional data sources, such as activity measurement for example through a sensor. This is supported in the artifact, as long as the data can be converted to FHIR observations. The goal of the artifact is to process and visualize the available data. As such, the implementation details of the data producers are not relevant, as long as the data is provided in the correct format.

The application should be tested and evaluated further in a practical setting to catch possible flaws that were not anticipated when evaluating the artifact on simulated data. The most important missing factors for practical use is authentication and authorization. There are various options for this, as discussed in section 4.7. The application is mostly ready to support multiple therapist, but mapping the users/principals to the database would need to be implemented along with authentication.

There is also potential to further improve the successful aspects of the artifact. Additional work to optimize scalability can be done. Some examples are given in section 4.6.2. Applying advanced algorithms to the flag functions, described in section 4.5.1, could also be beneficial. The application of pattern recognition or predictive analytics for this type of dashboard could be an interesting problem for further research. As discussed in section 2.2, there is a need for further structuring of the data from patient exercises before they can be used for analytics or visualization. Structuring and extracting data from these modules is an important topic for future research.

## References

- [1] G. V. Polanczyk, G. A. Salum, L. S. Sugaya, A. Caye, and L. A. Rohde, “Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents,” *Journal of Child Psychology and Psychiatry*, vol. 56, no. 3, pp. 345–365, 3 2015. [Online]. Available: <http://doi.wiley.com/10.1111/jcpp.12381>
- [2] H. Wittchen, F. Jacobi, J. Rehm, A. Gustavsson, M. Svensson, B. Jönsson, J. Olesen, C. Allgulander, J. Alonso, C. Faravelli, L. Fratiglioni, P. Jenum, R. Lieb, A. Maercker, J. van Os, M. Preisig, L. Salvador-Carulla, R. Simon, and H.-C. Steinhausen, “The size and burden of mental disorders and other disorders of the brain in Europe 2010,” *European Neuropsychopharmacology*, vol. 21, no. 9, pp. 655–679, 9 2011. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0924977X11001726?via%3Dihub>
- [3] S. Trautmann, J. Rehm, and H. Wittchen, “The economic costs of mental disorders,” *EMBO reports*, vol. 17, no. 9, pp. 1245–1249, 9 2016. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/27491723>
- [4] “Utvärdering av vård vid depression och ångestsyndrom,” Tech. Rep. [Online]. Available: [www.socialstyrelsen.se](http://www.socialstyrelsen.se)
- [5] D. Vigo, G. Thornicroft, and R. Atun, “Estimating the true global burden of mental illness,” *The Lancet Psychiatry*, vol. 3, no. 2, pp. 171–178, 2 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2215036615005052>
- [6] A. C. Butler, J. E. Chapman, E. M. Forman, and A. T. Beck, “The empirical status of cognitive-behavioral therapy: A review of meta-analyses,” *Clinical Psychology Review*, vol. 26, no. 1, pp. 17–31, 1 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0272735805001005>
- [7] J. D. Edinger, W. K. Wohlgemuth, R. A. Radtke, G. R. Marsh, and R. E. Quillian, “Cognitive Behavioral Therapy for Treatment of Chronic Primary Insomnia,” *JAMA*, vol. 285, no. 14, p. 1856, 4 2001. [Online]. Available: <http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.285.14.1856>
- [8] J. K. Smith and R. F. Józefowicz, “Diagnosis and treatment of somatoform disorders.” *Neurology. Clinical practice*, vol. 2, no. 2, pp. 94–102, 6 2012. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/29443321>
- [9] S. G. Hofmann, A. Asnaani, I. J. J. Vonk, A. T. Sawyer, and A. Fang, “The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses,” *Cognitive Therapy and Research*, vol. 36, no. 5, pp. 427–440, 10 2012. [Online]. Available: <http://link.springer.com/10.1007/s10608-012-9476-1>
- [10] R. Kohn, S. Saxena, I. Levay, and B. Saraceno, “The treatment gap in mental health care,” *Bulletin of the World Health Organization*, vol. 82, pp. 858–866, 2004. [Online]. Available: [https://www.scielosp.org/scielo.php?pid=S0042-96862004001100011&script=sci\\_arttext&tlng=en](https://www.scielosp.org/scielo.php?pid=S0042-96862004001100011&script=sci_arttext&tlng=en)

- [11] T. Nordgreen, R. Gjestad, G. Andersson, P. Carlbring, and O. E. Havik, “The effectiveness of guided internet-based cognitive behavioral therapy for social anxiety disorder in a routine care setting,” *Internet Interventions*, vol. 13, pp. 24–29, 9 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2214782918300204>
- [12] A. P. Folker, K. Mathiasen, S. M. Lauridsen, E. Stenderup, E. Dozeman, and M. P. Folker, “Implementing internet-delivered cognitive behavior therapy for common mental health disorders: A comparative case study of implementation challenges perceived by therapists and managers in five European internet services,” *Internet Interventions*, vol. 11, pp. 60–70, 3 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2214782917301203>
- [13] “Triage—Definition of Triage by Merriam-Webster.” [Online]. Available: <https://www.merriam-webster.com/dictionary/triage>
- [14] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A Design Science Research Methodology for Information Systems Research,” *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 12 2007. [Online]. Available: <http://www.tandfonline.com/doi/full/10.2753/MIS0742-1222240302>
- [15] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design Science in Information Systems Research,” Tech. Rep. 1, 2004. [Online]. Available: [https://wise.vub.ac.be/sites/default/files/thesis\\_info/design\\_science.pdf](https://wise.vub.ac.be/sites/default/files/thesis_info/design_science.pdf)
- [16] “eMeistring.” [Online]. Available: <https://helse-bergen.no/emeistring/>
- [17] “Cognitive Behavioral Therapy — Definition of Cognitive Behavioral Therapy by Merriam-Webster.” [Online]. Available: [https://www.merriam-webster.com/dictionary/cognitivebehavioraltherapy?utm\\_campaign=sd&utm\\_medium=serp&utm\\_source=jsonld](https://www.merriam-webster.com/dictionary/cognitivebehavioraltherapy?utm_campaign=sd&utm_medium=serp&utm_source=jsonld)
- [18] “What Is Cognitive Behavioral Therapy?” [Online]. Available: <https://www.apa.org/ptsd-guideline/patients-and-families/cognitive-behavioral>
- [19] A. D. Williams and G. Andrews, “The Effectiveness of Internet Cognitive Behavioural Therapy (iCBT) for Depression in Primary Care: A Quality Assurance Study,” *PLoS ONE*, vol. 8, no. 2, p. e57447, 2 2013. [Online]. Available: <https://dx.plos.org/10.1371/journal.pone.0057447>
- [20] S. A. Montgomery and M. Asberg, “A new depression scale designed to be sensitive to change,” *The British Journal of Psychiatry*, vol. 134, no. 4, pp. 382–389, 4 1979. [Online]. Available: <http://bjp.rcpsych.org/cgi/doi/10.1192/bjp.134.4.382>
- [21] K. Kroenke, R. L. Spitzer, and J. B. Williams, “The PHQ-9: validity of a brief depression severity measure.” *Journal of general internal medicine*, vol. 16, no. 9, pp. 606–13, 9 2001. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/11556941>

- [22] A. T. Beck, N. Epstein, G. Brown, and R. A. Steer, “An inventory for measuring clinical anxiety: Psychometric properties.” *Journal of Consulting and Clinical Psychology*, vol. 56, no. 6, pp. 893–897, 1988. [Online]. Available: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-006X.56.6.893>
- [23] “Montgomery-Åsberg Depression Rating Scale (MADRS).” [Online]. Available: <http://narr.bmap.ucla.edu/docs/MADRS.pdf>
- [24] W. van Ballegooijen, P. Cuijpers, A. van Straten, E. Karyotaki, G. Andersson, J. H. Smit, and H. Riper, “Adherence to Internet-Based and Face-to-Face Cognitive Behavioural Therapy for Depression: A Meta-Analysis,” *PLoS ONE*, vol. 9, no. 7, p. e100674, 7 2014. [Online]. Available: <https://dx.plos.org/10.1371/journal.pone.0100674>
- [25] M. Wilhelmsen, K. Lillevoll, M. B. Risør, R. Høifødt, M.-L. Johansen, K. Waterloo, M. Eisemann, and N. Kolstrup, “Motivation to persist with internet-based cognitive behavioural treatment using blended care: a qualitative study,” *BMC Psychiatry*, vol. 13, no. 1, p. 296, 12 2013. [Online]. Available: <http://bmcp psychiatry.biomedcentral.com/articles/10.1186/1471-244X-13-296>
- [26] J. Pleva and T. D. Wade, “Guided self-help versus pure self-help for perfectionism: A randomised controlled trial,” *Behaviour Research and Therapy*, vol. 45, no. 5, pp. 849–861, 5 2007. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0005796706001793>
- [27] K. L. Loeb, G. Wilson, J. S. Gilbert, and E. Labouvie, “Guided and unguided self-help for binge eating,” *Behaviour Research and Therapy*, vol. 38, no. 3, pp. 259–272, 3 2000. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0005796799000418>
- [28] B. F. Dear, M. Gandy, E. Karin, R. Fogliati, V. J. Fogliati, L. G. Staples, B. M. Wootton, L. Sharpe, and N. Titov, “The Pain Course: 12- and 24-Month Outcomes From a Randomized Controlled Trial of an Internet-Delivered Pain Management Program Provided With Different Levels of Clinician Support.” *The journal of pain : official journal of the American Pain Society*, vol. 19, no. 12, pp. 1491–1503, 12 2018. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/30099209>
- [29] T. Berger, F. Caspar, R. Richardson, B. Kneubühler, D. Sutter, and G. Andersson, “Internet-based treatment of social phobia: A randomized controlled trial comparing unguided with two types of guided self-help,” *Behaviour Research and Therapy*, vol. 49, no. 3, pp. 158–169, 3 2011. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0005796710002561>
- [30] G. Andersson, P. Cuijpers, P. Carlbring, H. Riper, and E. Hedman, “Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis.” *World psychiatry : official journal of the World Psychiatric Association (WPA)*, vol. 13, no. 3, pp. 288–95, 10 2014. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/25273302>

- [31] J. Bergheim, *Strukturering av data for internettbehandling av pasienter med depresjon (MSc thesis)*, 2018.
- [32] N. Titov, B. Dear, O. Nielssen, L. Staples, H. Hadjistavropoulos, M. Nugent, K. Adlam, T. Nordgreen, K. H. Bruvik, A. Hovland, A. Repål, K. Mathiasen, M. Kraepelien, K. Blom, C. Svanborg, N. Lindefors, and V. Kaldo, “ICBT in routine care: A descriptive analysis of successful clinics in five countries,” *Internet Interventions*, vol. 13, pp. 108–115, 9 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S221478291830037X>
- [33] A. for Healthcare Research, “Using Health IT in Practice Re-design: Impact of Health IT on Workflow: Health IT-Enabled Care Coordination and Redesign in Tennessee,” Tech. Rep. [Online]. Available: <https://www.healthit.gov/sites/default/files/playbook/pdf/hit-enabled-care-coordination-and-redesign-in-tn.pdf>
- [34] S. P. Bhavnani and A. M. Sitapati, “Virtual Care 2.0a Vision for the Future of Data-Driven Technology-Enabled Healthcare,” *Current Treatment Options in Cardiovascular Medicine*, vol. 21, no. 5, p. 21, 5 2019. [Online]. Available: <http://link.springer.com/10.1007/s11936-019-0727-2>
- [35] F. Holländare, G. Andersson, and I. Engström, “A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients.” *Journal of medical Internet research*, vol. 12, no. 5, p. e49, 12 2010. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21169165>
- [36] C. K. Thomas James, *Illuminating the Path: Research and Development Agenda for Visual Analytics.*, 2005. [Online]. Available: [https://ils.unc.edu/courses/2017\\_fall/inls641\\_001/books/RD\\_Agenda\\_VisualAnalytics.pdf](https://ils.unc.edu/courses/2017_fall/inls641_001/books/RD_Agenda_VisualAnalytics.pdf)
- [37] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, “Visual Analytics: Definition, Process, and Challenges,” in *Information Visualization*, A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, Eds. Berlin, Heidelberg: Springer-Verlag, 2008, ch. Visual Ana, pp. 154–175. [Online]. Available: [http://dx.doi.org/10.1007/978-3-540-70956-5\\_7](http://dx.doi.org/10.1007/978-3-540-70956-5_7)
- [38] B. A. Wilbanks and P. A. Langford, “A review of dashboards for data analytics in nursing,” pp. 545–549, 11 2014. [Online]. Available: <http://content.wkhealth.com/linkback/openurl%5C?sid=WKPTLP:landingpage%5C&an=00024665-201411000-00006>
- [39] M. Ghazisaeidi, R. Safdari, M. Torabi, M. Mirzaee, J. Farzi, and A. Goodini, “Development of Performance Dashboards in Healthcare Sector: Key Practical Issues.” *Acta informatica medica : AIM : journal of the Society for Medical Informatics of Bosnia & Herzegovina : casopis Drustva za medicinsku informatiku BiH*, vol. 23, no. 5, pp. 317–21, 10 2015. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/26635442>



- [40] E. Kroch, T. Vaughn, M. Koepke, S. Roman, D. Foster, S. Sinha, and S. Levey, "Hospital Boards and Quality Dashboards," *Journal Of Patient Safety*, vol. 2, no. 1, pp. 10–19, 2006. [Online]. Available: <https://oc.e.ovid.com/article/01209203-200603000-00003/HTML>
- [41] D. Dowding, R. Randell, P. Gardner, G. Fitzpatrick, P. Dykes, J. Favela, S. Hamer, Z. Whitewood-Moores, N. Hardiker, E. Borycki, and L. Currie, "Dashboards for improving patient care: Review of the literature," *International Journal of Medical Informatics*, vol. 84, no. 2, pp. 87–100, 2 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1386505614001890>
- [42] S. Negash and P. Gray, "CHAPTER 45 Business Intelligence," in *Handbook on Decision Support Systems 2: Variations*, 2008, pp. 175–193. [Online]. Available: [https://link.springer.com/content/pdf/10.1007%2F978-3-540-48716-6\\_9.pdf](https://link.springer.com/content/pdf/10.1007%2F978-3-540-48716-6_9.pdf)
- [43] J. G. Kang and K. H. Han, "A Business Activity Monitoring System Supporting Real-Time Business Performance Management," in *2008 Third International Conference on Convergence and Hybrid Information Technology*. IEEE, 11 2008, pp. 473–478. [Online]. Available: <http://ieeexplore.ieee.org/document/4682072/>
- [44] S. S. Khairat, A. Dukkipati, H. A. Lauria, T. Bice, D. Travers, and S. S. Carson, "The Impact of Visualization Dashboards on Quality of Care and Clinician Satisfaction: Integrative Literature Review." *JMIR human factors*, vol. 5, no. 2, p. e22, 5 2018. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/29853440>
- [45] A. Franklin, S. Gantela, S. Shifarrow, T. R. Johnson, D. J. Robinson, B. R. King, A. M. Mehta, C. L. Maddow, N. R. Hoot, V. Nguyen, A. Rubio, J. Zhang, and N. G. Okafor, "Dashboard visualizations: Supporting real-time throughput decision-making," *Journal of Biomedical Informatics*, vol. 71, pp. 211–221, 7 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1532046417301235?via%3Dihub>
- [46] W. Horn, C. Popow, and L. Unterasinger, "Metaphor graphics to visualize ICU data over time," *Intelligent Data Analysis in Medicine and Pharmacology IDAMAP98*, pp. 2–7, 1998.
- [47] W. G. Cole and J. G. Stewart, "Metaphor graphics to support integrated decision making with respiratory data," *International Journal of Clinical Monitoring and Computing*, vol. 10, no. 2, pp. 91–100, 5 1993. [Online]. Available: <http://link.springer.com/10.1007/BF01142279>
- [48] B. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," in *The Craft of Information Visualization*, 2003, pp. 364–371. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/B9781558609150500469>
- [49] "FHIR - Summary." [Online]. Available: <https://www.hl7.org/fhir/summary.html>

- [50] “Medical Device Interoperability — FDA.” [Online]. Available: <https://www.fda.gov/medical-devices/digital-health/medical-device-interoperability>
- [51] “Recommendation for using HL7 FHIR for data sharing,” 2019. [Online]. Available: <https://ehelse.no/standarder/ikke-standarder/recommendation-for-using-hl7-fhir-for-data-sharing>
- [52] “Publicly Available FHIR Servers for testing - HL7Wiki.” [Online]. Available: [http://wiki.hl7.org/index.php?title=Publicly\\_Available\\_FHIR\\_Servers\\_for\\_testing](http://wiki.hl7.org/index.php?title=Publicly_Available_FHIR_Servers_for_testing)
- [53] HL7, “JPA Server - HAPI FHIR.” [Online]. Available: [http://hapifhir.io/doc\\_jpa.html](http://hapifhir.io/doc_jpa.html)
- [54] “Search - Page Count - FHIR v4.0.0.” [Online]. Available: <https://www.hl7.org/fhir/search.html#count>
- [55] “Security - FHIR v4.0.0.” [Online]. Available: <https://www.hl7.org/fhir/security.html>
- [56] M. L. Braunstein and M. L. Braunstein, “SMART on FHIR,” 2018. [Online]. Available: <http://docs.smarthealthit.org/>
- [57] C. U. Lehmann, P. O. Rozo, and K. Holmes, “Practical Visualization of Multivariate Time Series Data in a Neonatal ICU,” *Vahc*, pp. 1–4, 2010. [Online]. Available: <http://ccom.uprrp.edu/~pordonez/VAHC2010-Lehmann.pdf>
- [58] R. Brendbekken, A. Harris, H. Ursin, H. R. Eriksen, and T. Tangen, “Multidisciplinary Intervention in Patients with Musculoskeletal Pain: a Randomized Clinical Trial.” *International journal of behavioral medicine*, vol. 23, no. 1, pp. 1–11, 2 2016. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/25920707>
- [59] A. Faiola, P. Srinivas, and J. Duke, “Supporting Clinical Cognition: A Human-Centered Approach to a Novel ICU Information Visualization Dashboard.” *AMIA ... Annual Symposium proceedings. AMIA Symposium*, vol. 2015, pp. 560–9, 2015. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/26958190>
- [60] A. Dagliati, L. Sacchi, V. Tibollo, G. Cogni, M. Teliti, A. Martinez-Millana, V. Traver, D. Segagni, J. Posada, M. Ottaviano, G. Fico, M. T. Arredondo, P. De Cata, L. Chiovato, and R. Bellazzi, “A dashboard-based system for supporting diabetes care,” *Journal of the American Medical Informatics Association*, vol. 25, no. 5, pp. 538–547, 5 2018. [Online]. Available: <https://academic.oup.com/jamia/article/25/5/538/4835587>
- [61] NHI.no, “Rettledning ved bruk av MADRS.” [Online]. Available: <https://nhi.no/skjema-og-kalkulatorer/kalkulatorer/psykiatri/montgomery-and-aasberg-depression-rating-scale-madrs/>

- [62] J. Walonoski, M. Kramer, J. Nichols, A. Quina, C. Moesel, D. Hall, C. Duffett, K. Dube, T. Gallagher, and S. McLachlan, “Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record,” *Journal of the American Medical Informatics Association*, vol. 25, no. 3, pp. 230–238, 3 2018. [Online]. Available: <https://academic.oup.com/jamia/article/25/3/230/4098271>
- [63] J. Nielsen, “Enhancing the explanatory power of usability heuristics,” in *Conference companion on Human factors in computing systems - CHI '94*, 1994, p. 210. [Online]. Available: <http://cs.ashoka.edu.in/cs102/papers/Nielsen94-Enhancing-Heuristics.pdf>
- [64] “Memory Recognition and Recall in User Interfaces.” [Online]. Available: <https://www.nngroup.com/articles/recognition-and-recall/>
- [65] “D3.js - Radar Chart or Spider Chart - bl.ocks.org.” [Online]. Available: <http://bl.ocks.org/nbremer/6506614>
- [66] “Matplotlib: Python plotting Matplotlib 3.1.0 documentation.” [Online]. Available: <https://matplotlib.org/>
- [67] “Seaborn: statistical data visualization seaborn 0.9.0 documentation.” [Online]. Available: <https://seaborn.pydata.org/>
- [68] “Dash by Plotly - Plotly.” [Online]. Available: <https://plot.ly/products/dash/>
- [69] HL7, “FHIR C# client library.” [Online]. Available: <https://github.com/FirelyTeam/fhir-net-api>
- [70] “Smart on FHIR python client.” [Online]. Available: <https://github.com/smart-on-fhir/client-py>
- [71] “ASP.NET Identity - ASP.NET 4.x — Microsoft Docs.” [Online]. Available: <https://docs.microsoft.com/en-us/aspnet/identity/overview/getting-started/introduction-to-aspnet-identity>
- [72] “Entity Framework 6 — Microsoft Docs.” [Online]. Available: <https://docs.microsoft.com/en-us/ef/ef6/>
- [73] MongoDB, “C# and .NET MongoDB Driver MongoDB Ecosystem.” [Online]. Available: <https://docs.mongodb.com/ecosystem/drivers/csharp/>
- [74] “BSON (Binary JSON) Serialization.” [Online]. Available: <http://bsonspec.org/>
- [75] Highsoft AS, “Interactive JavaScript charts for your webpage — Highcharts,” 2016. [Online]. Available: <https://www.highcharts.com/highcharts.com>
- [76] M. Bostock, “D3.js - Data-Driven Documents,” 2009. [Online]. Available: <https://d3js.org/>

- [77] HL7, “Subscription - FHIR v4.0.0.” [Online]. Available: <https://www.hl7.org/fhir/subscription.html>
- [78] “Graphql - FHIR v4.1.0.” [Online]. Available: <https://build.fhir.org/graphql.html>
- [79] S. Hove and B. Anda, “Experiences from Conducting Semi-structured Interviews in Empirical Software Engineering Research,” in *11th IEEE International Software Metrics Symposium (METRICS'05)*. IEEE, 2005, pp. 23–23. [Online]. Available: <http://ieeexplore.ieee.org/document/1509301/>
- [80] “Stopwatch Class (System.Diagnostics) — Microsoft Docs.” [Online]. Available: <https://docs.microsoft.com/en-us/dotnet/api/system.diagnostics.stopwatch?view=netframework-4.8>
- [81] “Apache Tomcat® - Welcome!” [Online]. Available: <http://tomcat.apache.org/>
- [82] “Paged results SimplifierDocs 3.5.1 documentation.” [Online]. Available: <http://docs.simplifier.net/fhirnetapi/client/paging.html>
- [83] “Grid system · Bootstrap.” [Online]. Available: <https://getbootstrap.com/docs/4.0/layout/grid/>
- [84] P. J. Lukyanenko Roman, Evermann Joerg, “Instantiation Validity in IS Design Research,” *Advancing the Impact of Design Science: Moving from Theory to Practice*, 2014. [Online]. Available: <https://link.springer.com/content/pdf/10.1007%2F978-3-319-06701-8.pdf>
- [85] “MongoDB - Convention for ignoring all empty collection types in serialization?” [Online]. Available: <https://jira.mongodb.org/browse/CSHARP-765>
- [86] “Plotly.js — JavaScript Graphing Library.” [Online]. Available: <https://plot.ly/javascript/>

## Appendix A

### Answers from semi-structured interviews

The answers to the questions in the semi-structured interview for evaluating the artifact is included here. As the full transcriptions are both lengthy and hard to read, the answers are summarized as points. Because the interviews were performed in Norwegian, both questions and answers are translated. The answers are further paraphrased for brevity, and grouped by theme for readability. The method for conducting the interviews are described in section 5.1.

Q1: Could you say something about your experience with iCBT?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Worked full time in psychiatry since 2003, part time before.</li> <li>• Further education within psychodynamic psychotherapy, CBT and iCBT.</li> <li>• Worked 5 years in eMeistring, combined with having regular patients in polyclinic.</li> </ul>	<ul style="list-style-type: none"> <li>• Education within cognitive sports-psychology.</li> <li>• Did some studies on CBT on back-patients and patients with exhaustion syndrome.</li> <li>• No previous experience with iCBT, currently in the process of designing a self-help program.</li> </ul>

Q2: Which types of patient data do you normally consider?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Varies between patients.</li> <li>• Initial overview through conversation with the patient.</li> <li>• Quantitative measurements like MADRS are useful for some situations, for example depression.</li> <li>• Visual observations and physiological measurements are also used.</li> </ul>	<ul style="list-style-type: none"> <li>• Questionnaires.</li> <li>• Physiological measurements like sleep, heart-frequency, blood tests.</li> <li>• Register data such as absence due to illness, diagnoses, and specific findings.</li> </ul>

Q3: Which types of data would you like to see in the tool?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Patients life events.</li> <li>• What patients currently are doing.</li> <li>• More details for serious events. For example if a patient reports trouble with appetite and sleep, it would be useful to get further descriptions on these problems.</li> <li>• Messages from the patients.</li> <li>• The content of the treatment modules.</li> <li>• Metadata about the modules such as which module is a patient in and what they have delivered.</li> </ul>	<ul style="list-style-type: none"> <li>• Information about other treatments the patient is receiving.</li> <li>• More data of the same types displayed in the tool.</li> <li>• Activity measurements. For example if patients are going for walks or training.</li> </ul>

Q4: Which tools have you (or other therapists) used to process patient-data?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Simple bar charts to display mood.</li> <li>• The bars display the values within a single questionnaire.</li> <li>• Can only display a single questionnaire at a time.</li> <li>• Have to click multiple times to get into a new questionnaire.</li> </ul>	<ul style="list-style-type: none"> <li>• Refers to a PhD student using spider charts as a communication tool towards the patient (see section 2.5.2).</li> <li>• Some therapists read the medical journals, some read the "raw" patient-filled questionnaires like BDI.</li> <li>• A lot of therapists like to only work on paper.</li> </ul>

Q5: Which features of the tool is applicable for your situation?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Time spent in program.</li> <li>• Flags can be useful.</li> <li>• Progression.</li> <li>• Filtering the table is useful.</li> </ul>	<ul style="list-style-type: none"> <li>• Warnings are useful. The list over warnings should preferably be short.</li> <li>• Flags are essential. Yellow and red flags are suggested.</li> <li>• Progression: improving, declining and steady.</li> <li>• Urgency score, possibly based on other measurements than MADRS.</li> <li>• Spider chart displaying changes. This can be useful for patient communication.</li> </ul>

Q6: Which new features would you like to see in the tool?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Comparing patients towards each other.</li> <li>• Compare patients within filtered groups.</li> <li>• Zoom in on line charts, select a subset of the time period to display.</li> <li>• Which module a patient is in.</li> </ul>	<ul style="list-style-type: none"> <li>• Option to start the scale of line chart y-axes at 0.</li> <li>• Option for showing the measurements on the same time axis, for comparison between measurements.</li> <li>• Option to combine measurements in a single line chart, for comparison.</li> <li>• Line charts showing ratios between observations.</li> <li>• Red and yellow flags.</li> <li>• Option for showing multiple spider charts on the same screen, for multiple questionnaires.</li> <li>• Possibility to select a subset of patients, and for the pie chart to update to the selected patients.</li> <li>• Export data out of the application, for example export to a csv file.</li> </ul>

Q7: How do you think this kind of tool would fit in therapists current workflow?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• It would be more useful than what the therapists currently use.</li> <li>• If all the necessary information was presented, it would fit better than the existing solution.</li> </ul>	<ul style="list-style-type: none"> <li>• Therapists in CBT could have looked at it before and/or after a consultation.</li> <li>• Would possibly be useful for a lot of clinicians, although a lot of clinicians still use paper primarily.</li> <li>• Would have been more useful with line charts on the same time axis.</li> </ul>

Q8: What do you think about the visualizations provided in the tool?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Good contrasts between the colors in the pie chart.</li> <li>• Displaying weeks instead of months in treatment would be better.</li> <li>• Line chart with the sum of MADRS is useful.</li> <li>• The spider chart is useful to get a quick overview of the patient.</li> </ul>	<ul style="list-style-type: none"> <li>• Line charts and spider charts are appropriate for these data.</li> </ul>

Q9: Which of the visualizations in the tool do you think can help therapists make decisions?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• The pie chart is useful to get an overview. This can be beneficial for logistics, such as planning capacity for new patients.</li> <li>• The line chart over MADRS good for seeing patients symptoms. If a patient spikes upwards, it signals that the patient needs to be examined further.</li> </ul>	<ul style="list-style-type: none"> <li>• The flags are probably the most important.</li> <li>• Progression, depending on how it is calculated.</li> <li>• Line charts summing over the questionnaires are useful to see if there is improvement, also over a longer time span.</li> </ul>



Q10: Which methods would you use to determine the urgency of treatment for a patient?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Will look at all the patients if there is sufficient time.</li> <li>• Alarms are used in current practice, for example if a patient is too depressed.</li> <li>• Consider the amount of messages. A large amount of messages indicates the patient wants to communicate something.</li> </ul>	

Q11: Which status variables would you like to see for each patient?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Name.</li> <li>• How far the patient is in the program.</li> <li>• Time of starting the program.</li> <li>• Time of ending the program.</li> <li>• Diagnose/which problems.</li> <li>• Which treatment program.</li> </ul>	<ul style="list-style-type: none"> <li>• Time of last consultation.</li> <li>• Time of last contact between therapist and patient.</li> <li>• Being able to choose which status variables are displayed would be beneficial. This could be done similarly to how observations are selected in the Detail View.</li> </ul>

Q12: Which information would you like to see for the patient population?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• A table over the patients, with useful variables.</li> <li>• Possibility for filtering the table based on the variables.</li> <li>• Dividing the patients by progression is useful, for example presented in the pie chart.</li> </ul>	<ul style="list-style-type: none"> <li>• Why patients are improving/declining.</li> <li>• Means for various measurements across all the patients, for example the questionnaires.</li> </ul>

Q15: Are there other kinds of functionality for decision support you would like to see in a tool like this?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Combining functionality. For example combine "time spent in program" with progression/module to see who is taking too long.</li> <li>• Color marking in combination with filtering. For example: Change the color of patient names when the bar chart is clicked and the matching patients are placed on the top.</li> </ul>	<ul style="list-style-type: none"> <li>• Suggests option for the user to supply their own data.</li> </ul>

Q14: How much time would you estimate therapist in iCBT spend processing patient data?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• If we don't consider paperwork, such as writing, at least half of the time is spent getting an overview of what has happened.</li> </ul>	

Q15: Which benefits can be achieved by having an overview of the patient population?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• It is useful to know who/what to prioritize.</li> <li>• Logistics. When patients finish the treatment, the therapist will have capacity for new patients.</li> <li>• It is useful to compare with other therapists. For example look for differences, and how these can be explained.</li> <li>• Comparison with other therapists can also help to determine how to balance the workload across patients.</li> </ul>	<ul style="list-style-type: none"> <li>• If a lot of patients are declining, this is a red flag.</li> </ul>

Q16: Do you have any thoughts about how to reduce the amount of time therapists spend processing data?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• Fast processes and good overview would help.</li> <li>• Not having to spend time on unproductive activities.</li> </ul>	<ul style="list-style-type: none"> <li>• It is important that patients can fill in the required data, without the therapist needing to spend time on it.</li> <li>• Fast overview of a patient would save a lot of time. Seeing data on the same time axis would help for this.</li> </ul>

Q17: Do you have other thoughts about how to scale the iCBT solution to enable more patients per therapist?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• eMeistring is an attempt at reducing time spent on unproductive activities, such as travelling or talking about things without effect.</li> <li>• If the goal is to further reduce the time, the question is how little time you need per patient to get the effect of guided iCBT.</li> <li>• Ideally everything could be automated without losing the effect of guidance, but unguided self-help has shown problems with patients not wanting to put in the effort.</li> <li>• To increase therapist effectivity, they will need to spend less time processing information.</li> <li>• Considering data, overview is what is important. Pre-made procedures could also help.</li> <li>• Time can be saved on reading and writing.</li> </ul>	<ul style="list-style-type: none"> <li>• Mentions a colleague organizing group CBT over the internet, with video.</li> <li>• Self-help programs.</li> </ul>

Q18: Do you have any other comments towards the functionality presented in the tool?	
Interviewee 1	Interviewee 2
<ul style="list-style-type: none"> <li>• The functionality is an improvement to the current situation.</li> <li>• If it would be generalized to all types of data they need in eMeistring, it would be very useful.</li> <li>• There are multiple other data sources than questionnaires that are relevant.</li> </ul>	<ul style="list-style-type: none"> <li>• It is not only useful for eMeistring. It could, for example, potentially be applied to restitution among athletes.</li> <li>• It could be useful to see patterns that are not necessarily clinical in nature.</li> <li>• Subjectively most interested in the information for the individual patients.</li> <li>• It would be interesting to see the tool applied on with long treatment periods.</li> </ul>

## Appendix B

### Source code

The source code for the design science artifact is open source and available at:  
<https://github.com/NikolaiGrieg/TherapyDashboard>.