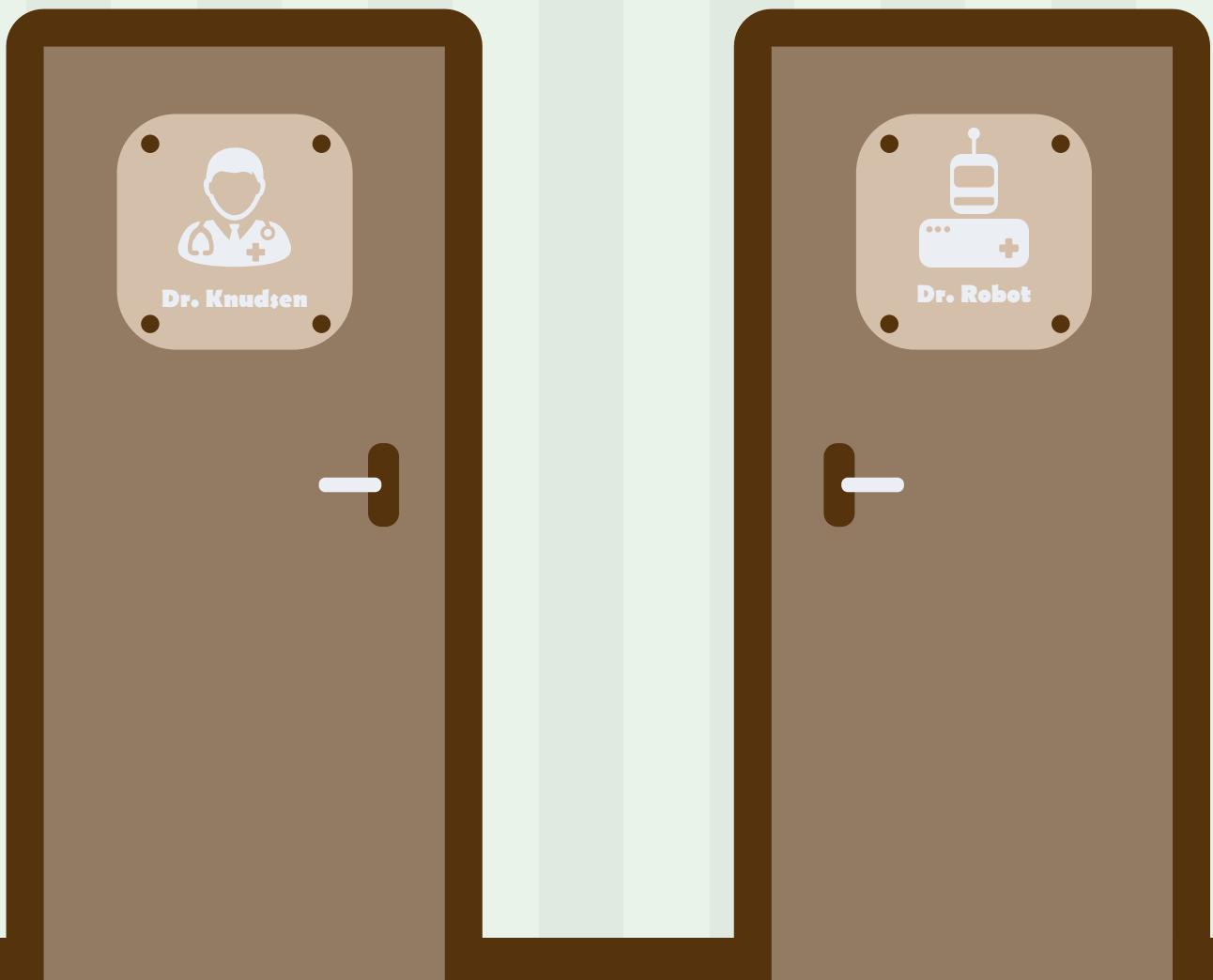


Optimising User Trust in Decision Aids in the Healthcare Domain

A National Decision Aid for Medication Prescription

June 2019



Title: Optimising User Trust in Decision Aids in the Healthcare Domain
Theme: Master's Thesis Report
Case: A national decision aid for medication prescription

Number of pages: 52
Number of appendix pages: 83

Semester: 10th semester
Project period: 1st of February 2019 to 6th of June 2019

University supervisor: Dorte Hammershøi

Authors: Mikkel Bay Revsbech & Morten Adelsen Jakobsen
Group: 19gr1087

Synopsis

To explore healthcare professionals' trust in decision aids and other healthcare professionals as well as which factors affect these perceptions, an experimental pairwise comparison study with 28 healthcare professionals and medicine students is conducted. The study is based on 10 perceptual statements about two adviser characters (a consultant and a decision aid) that were found, developed, and described based on semi-structured and contextual interviews with healthcare professionals and professors specialising in the Danish healthcare sector.

The experimental study shows that healthcare professionals perceive the statements to be true for the decision aid to a larger extent than the consultant for 3 statements. Healthcare professionals have likewise perceived the statements to be true for the consultant to a larger extent than the decision aid in 5 statements. The statements are not likely to be significantly different in 2 statements.

Based on the combined results of the exploratory and experimental studies it is concluded that healthcare professionals are not averse to advice from decision aids, according to the three factors perceived as more true for the decision aid. Furthermore, it is found that the decision aid can be improved on the remaining 8 factors to perform as well as or better than the consultant. Changes to the design of the decision aid are suggested.

The mixed method approach to explore people's perception of algorithms and algorithmic advice in the form of decision aids is found to provide useful results. The results provide insights into factors that affect healthcare professionals' perception of advice from different advisers. The results can be turned into useful improvements and considerations for the design process of decision aids and algorithmic advice.

Resumé

Algoritmer og kunstig intelligens overtager store dele af folks hverdag, og er i forvejen til stede i blandt andet smartphones, biler, hospitaler og banker. Baseret på de store mængder data, der hver dag indsamles gennem sensor teknologier, er det muligt at påvirke og hjælpe mennesker i deres beslutningstagning. Teknologien har et stort potentiale til at effektivisere hele brancher, heriblandt sundhedssektoren, hvor der investeres i udvikling af digitale beslutningsstøttesystemer. For at udnytte algoritmers fulde potentiale kræver det tillid til anvendelsen af dem.

For at undersøge sundhedspersonales tillid til beslutningsstøttesystemer og andre sundhedspersonale, samt hvilke faktorer der påvirker deres opfattelser, gennemføres en eksperimentel parvis sammenligningsundersøgelse (*pairwise comparison*) på 28 sundhedsprofessionelle og medicinstuderende. Den eksperimentelle undersøgelse er baseret på 10 perceptuelle udtafelser om to rådgivningskarakterer – en overlæge og et beslutningsstøttesystem. Rådgivningskaraktererne er fundet, beskrevet og designet baseret på 3 semi-strukturelle interviews med professorer og projektledere inden for sundhedsvæsenet samt 1 kontekstuelt interview med en overlæge på kardiologisk afdeling på et sygehus i Region Nordjylland.

Den eksperimentelle undersøgelse viser, at sundhedspersonale opfatter udsagn til at være rigtige for beslutningsstøttesystemet i højere grad end overlægen for 3 udsagn. Ydermere opfatter sunhedspersonale udsagn til at være rigtige for overlægen i højere grad end beslutningsstøttesystemet for 5 udsagn. For 2 udsagn er det ikke sandsynligt at der er signifikant forskel for de to rådgivningskarakterer.

Baseret på de kombinerede resultater af det eksplorative og det eksperimentelle studie konkluderes det at sundhedspersonale generelt ikke er modvillige overfor rådgivning fra beslutningstøttesystemer. Ifølge de tre udsagn, der opfattes som mere rigtige for beslutningstøttesystemet i forhold til overlægen, findes det, at der er en stærk tiltro til, at beslutningstøttesystemet rådgiver baseret på videnskabelig evidens og kan omstille sig til ny viden i højere grad end overlægen. Desuden findes den grundlag for at forbedre beslutningsstøttesystemet med konkrete designforslag baseret på de 5 udsagn, der vurderes som mere rigtige for overlægen. Det drejer sig om at lægge større vægt på mængden af data, der tages i betragtning af beslutningsstøttesystemet, bedre visualiseringer og simuleringer af råd og konsekvenser af disse, at de mest relevante data altid er tilgængelige for at opfattelsen af at tage hensyn til patientens behov, basere rådgivning på et fuldstændigt billede af patienten og bedre forklare ræsonnementet for rådgivningen.

Denne mixed method tilgang til at undersøge sundhedspersonales opfattelse af algoritmer og råd fra disse i form af beslutningstøttesystemer, findes at være god til at levere brugbare resultater. Resultaterne giver indsigt i de faktorer, som påvirker sundhedspersonales opfattelse af råd fra forskellige rådgivere. Resultaterne kan omdannes til brugbare forbedringer og overvejelser til designprocessen og, hvordan beslutningsstøttesystemer designes for at optimere tilliden til de råd der gives.

Foreword

This master's thesis is carried out in the period 1st of February to 6th of June by Mikkel Bay and Morten Adelsen Jakobsen (group 1087) as the closing dissertation on the masters degree programme in Engineering Psychology at the institute of Electronic Systems at Aalborg University. The thesis supervisor is Dorte Hammershøi, professor at the institute of Electronic Systems. The thesis authors would like to thank the participating subjects of the experimental study as well as the interview study participants for taking time out of their schedule to participate in both interviews, pilot testing, review of experimental design as well as the pairwise experiment. As part of the interview process participants were given the option of confidentiality as well as to go through the interview recording and transcript before final hand-in. No participants, however expressed the desire to do so. Therefore, the transcriptions and following analysis of the interviews are the authors' interpretation of the recorded interview and should be treated as such.

Reading instructions

This report is divided into chapters with a short introduction to each chapter explaining the content of the chapter, references to chapters and sections are written as the following example: [chapter 1](#) referencing chapter 1 and [section 4.4](#) referencing section 4 in chapter 4.

To reference figures and tables each is numbered according to the chapter in which it appears and the order in which it appears in the given chapter. Referencing a figure looks like the following: [Figure 5.5](#), meaning a reference to figure number 5 in chapter 5. A table is referenced as: [Table 5.4](#), meaning a reference to table number 3 in chapter 5.

External references are written using the APA referencing system. Hence, references are written as (Last name(s), year). E.g. ([Harari, 2016](#)). In-text references are written as Last name(s) (year). E.g. as used in the reference [Harari \(2016\)](#). All references can be found alphabetically ordered in the bibliography in the back of the report.

Appendices referenced in the report can be found either in the back of the report or as a separate appendix report, in which appendices are assigned a letter e.g. see [Appendix A](#). Furthermore, some appendices are only found as part of a zip-file and are referenced as electronic appendices. These are referred to with a reference to the list of electronic appendices provided in [Appendix L](#). These are found in the attached file: ElectronicAppendix.zip. All references made in this report function as links that takes the reader to the reference which is clicked except for the electronic appendices which as mentioned is found in the attached file.

Table of content

Chapter 1 Introduction	1
1.1 The rise of sensor technologies and big data	1
1.2 Algorithms, machine learning and artificial intelligence	1
Chapter 2 Problem analysis	3
2.1 People's trust in algorithmic advice	3
Chapter 3 Research questions	5
Chapter 4 Exploratory study	7
4.1 Preparation for the experimental study	7
4.2 Considerations about methods	7
4.3 Exploring people and decision aids in the healthcare domain	8
4.4 Results and analysis	11
Chapter 5 Experimental Study	23
5.1 Study description	23
5.2 Method considerations	23
5.3 Experimental design	26
5.4 Results and data analysis	30
Chapter 6 General discussion	39
6.1 Healthcare professionals and trust in decision aids	39
6.2 Trust in advice from other healthcare professionals	40
6.3 Perception of advice from a consultant and a decision aid	41
6.4 Improving decision aids based on probabilistic choice models	42
6.5 Evaluation of chosen methods	45
6.6 Future work	47
6.7 Conclusion	47
References	49
Appendix	1
Appendix A Interview I	3
A.1 Interview Guide	3
A.2 Transcription	5
A.3 Analysis	13
Appendix B Interview II	23
B.1 Interview Guide	23
B.2 Transcription	25
B.3 Analysis	31
Appendix C Interview III	43
C.1 Interview Guide	43
C.2 Transcription	44
C.3 Analysis	48
Appendix D Contextual interview notes	55
Appendix E Medical Information	57
Appendix F Review session	59
F.1 Session design	59

F.2 Results and changes	60
Appendix G Consultant	61
Appendix H PrescriptAid	63
Appendix I Pairwise software description	65
I.1 Preparation and planning	65
I.2 The code	65
I.3 User interface and functionality	73
Appendix J Pilot experiment	77
J.1 Results	78
Appendix K Script	81
Appendix L Electronic appendix overview	83
L.1 PairwiseComparison	83
L.2 RawData	83
L.3 ExitInterview	83
L.4 Rscript	83

Introduction

This chapter functions as an introduction to the technological developments concerning algorithmic and sensor technologies that based on large amounts of data, big data, and through machine learning and artificial intelligence, has the possibility to drastically change how we as humans make decisions and which decisions are made by humans.

1.1 The rise of sensor technologies and big data

During the 21st century the amount of sensor technologies in every day people's lives have risen drastically (Ernst, Merola, & Samaan, 2018). Most people in level 1 and 2 income countries willingly carry around their cell phone, laptop, or smartwatch containing sensors including accelerometer, gyroscope, magnetometer, GPS, proximity sensors, microphones, touch screen sensors, and finger print sensors. These sensors detect or measure physical properties in real time and provide corresponding outputs or measurements in response.

For most people, sensors collecting data has become a part of everyday life. When surfing the internet, most websites log keystrokes, mouse-clicks, and typed in commands such as searches. When moving around public places such as airports metal detectors, cameras and the likes are used for security. Furthermore, sensors are widely used when borrowing books on libraries, when going to the hospital or used by health and fitness technology to measure performance and progress.

During the last decade an increased amount of connectivity and internet access has led to more and more data being collected and sent to databases to be saved (Ernst et al., 2018). With the large amount of connected sensors comes an even larger amount of collected measurements and data from these sensors. These very large data sets that may be analysed by computers to reveal patterns, trends, and associations are called big data (Oxford Dictionaries, 2019).

1.2 Algorithms, machine learning and artificial intelligence

Analysing and revealing patterns in large amounts of data, big data, makes it possible to create algorithms based on the revealed patterns or even use machine learning to create the algorithms directly based on the collected data. Machine learning makes it possible to create adaptive algorithms that has the potential of self-learning and adaptation based on new data presented to them.

The enhancement of decision aids, machine learning, artificial intelligence, and algorithms in general is according to the International Labour Organization (ILO) of the United Nations (UN) going to take over and reshape human jobs faster and more effectively than ever before (Ernst et al., 2018). According to Harari (2016) this technological development has the potential to create a useless class of humans that are not able to adapt to the changes of the job market and are hence replaced by algorithmic technologies.

Artificial intelligence technologies have already taken over large portions of our everyday lives. AI is present in phones, cars, matchmaking services, hospitals, banks, and all over media services. Digital assistants such as *Google Assistant*, *Amazon Alexa*, *Microsoft Cortana*, *Baidu Duer*, and *Apple's Siri* are ready to answer complex spoken questions, schedule appointments, and even call restaurants on our behalf through home appliances, phones, and laptops (Hoy, 2018).

Furthermore, companies and governments are developing new technologies such as self-driving vehicles (Aalborg Havn, 2018; Waymo, 2019) and glucose-monitoring contact lenses (Mendoza, 2014; Park et al., 2018), relying heavily on sensor technology. By measuring and responding to large amounts of physical properties these technologies has the potential to outperform humans in specialised fields. Algorithms already beat doctors and pathologists in predicting cancer survival (Haenssle et al., 2018), heart attack occurrence (Goldman et al., 1977), and disease severity (Einhorn, 1972).

In other cases these technologies are used as decision aids to support humans in their decision making process. Decision aids can range from simple statistical models to more complex machine learning based decision aids. These decision aids can range from more commonly known decision aids such as Google Maps, smartphone keyboards that suggest words, email software that writes emails based on previously written messages, and cars such as the Tesla which helps while driving to more specialised decision aids such as BEACON caresystem that based on machine learning helps healthcare professionals make decisions about highly sedated patients. The Danish ministry of health and the elderly has invested large sums of money in the development of new decision aids to the Danish healthcare sector regarding medicament ordination and monitoring of the elderly population and predictions from gathered data ([Sundheds- og Ældreministeriet, 2018a](#)).

Hence, it is found relevant and interesting to analyse the problems that arise when people have the choice to or are forced to trust algorithmic advice and decision aids or have to trust choices made by algorithms and machine learning technologies, and what affects their trust in algorithms. This gives the initial problem statement:

Do people trust decision aids and decisions made by algorithms? And what affects people's trust in algorithms?

This initial problem statement and the problems related to it are analysed in the following chapter.

2

Problem analysis

This chapter contains an analysis of problems regarding people's trust in decision aids, algorithms and algorithmic advice.

2.1 People's trust in algorithmic advice

It is a common held belief that humans distrust algorithmic outputs, especially in the decision making and judgement literature (Dietvorst, Simmons, & Massey, 2015; Logg, Minson, & Moore, 2019; Meehl, 1954). This distrust is sometimes referred to as algorithmic aversion (Dietvorst et al., 2015).

Not much research is made regarding the causes of algorithmic aversion. Some cited reasons for the cause of algorithm aversion are; the inability for algorithms to learn (Dawes, 1979), the ability for human forecasts to improve through experience (Highhouse, 2008), the notion that algorithms cannot consider individual targets (Grove & Meehl, 1996), the inability of algorithms to incorporate qualitative data (Grove & Meehl, 1996), a desire for forecasts to be perfect (Dawes, 1979; Einhorn, 1986; Highhouse, 2008), the notion that algorithms are dehumanising (Dawes, 1979; Grove & Meehl, 1996), the ethics in relying on algorithms with important decisions (Dawes, 1979; Shariff, Bonnefon, & Rahwan, 2017), overconfidence in own performance (Sieck & Arkes, 2005), too little or too much transparency (Koo et al., 2015), and not knowing the superiority of algorithms (Shariff et al., 2017). This research is, however, primarily based on anecdotal experience rather than empirical evidence. Hence, it is important to further research the anecdotal statements before making assumptions based on these data.

2.1.1 Experimental studies of trust in algorithmic advice

Empirical research have shown that seeing algorithms err makes people less confident in them (Dietvorst et al., 2015; Prahl & Van Swol, 2017) and that being able to (even slightly) modify the algorithm makes people more confident in them (Dietvorst, Simmons, & Massey, 2016). Some research shows that people often prefer forecasts made by humans rather than those made by superior algorithms (Diab, Pui, Yankelevich, & Highhouse, 2011; Eastwood, Snook, & Luther, 2012). It is shown that people tend to weigh human input more strongly than algorithmic input (Önkal, Goodwin, Thomson, Gönül, & Pollock, 2009; Promberger & Baron, 2006). Furthermore, it is shown that professionals who seek advice from algorithms are judged more harshly than those seeking advice from a human (Shaffer, Probst, Merkle, Arkes, & Medow, 2013; Wolf, 2014).

Despite research showing people's algorithmic aversion, new research challenges the assertion that people are averse to algorithms (Dietvorst et al., 2015; Logg et al., 2019; Prahl & Van Swol, 2017). It is shown that people readily rely on algorithmic advice before seeing the algorithm err. Furthermore, people choose algorithmic judgement over their own judgement (Logg et al., 2019). This willingness to trust algorithmic advice is coined algorithmic appreciation, and is displayed regardless of age. Algorithmic appreciation is however seen to be lower in less numerate people as well as among experts who were less likely to be open for taking any advice. Hence, resulting in lowered accuracy in their judgements.

The majority of these experimental studies are performed in laboratories or as remote questionnaire studies, which results in data with a low external and ecological validity, hence making it difficult to apply the results directly to the real world and to contexts outside of the experimental studies themselves. Some studies (Dietvorst et al., 2015; Logg et al., 2019; Prahl & Van Swol, 2017) do, however, show a tendency for humans to trust algorithmic advice to a larger extent than the advice of another person or a group of people in simple tasks such as guessing a persons weight or predicting how far a song will rank on the top lists. While other studies (Önkal et al., 2009; Promberger & Baron, 2006) show the opposite effect when humans choose between algorithmic and expert human advice in more complex issues such as choosing where to place money on the stock market or choosing a treatment for a disease.

2.1.2 Experienced professionals and trust in algorithms

According to Logg et al. (2019) expert professionals are found to be less likely to take advice from others, even superior algorithms, and thus have lower accuracy in their judgements. Hence, it is found interesting to research experts' apparent aversion to trust advice from algorithms. As mentioned, other research shows that people are more likely to trust human expert professionals than they are algorithms (Önkal et al., 2009; Promberger & Baron, 2006). This preference for human expert advice seems stronger than any algorithm appreciation found in the experimental studies. Based on these studies, it is found interesting to research whether human experts distrust decision aids and the generated advice. This is summarised in the following research question:

Do expert professionals distrust decision aids and algorithmic advice, and why?

Similarities between human experts and black box algorithms

In their study, Logg et al. (2019), mention the similarities between human experts and black box algorithms in their way of processing data. Black box algorithms are algorithms in which the input and the output are known but it is not possible to map how the given output was calculated exactly. This is often the case when it comes to algorithms and decision aids made using machine learning. This is comparable to the black box behaviour of advice and decisions made by humans and human experts. This leads to the following research question:

To what extent do expert professionals trust advice from other expert professionals, and why?

Algorithmic advice presentation

In situations where algorithms in the form of statistical models or artificial intelligence function as decision aids, the data needs to be presented in a way such that the optimal decisions are made while at the same time optimising the appreciation of the decision aids. Hence, it is important to focus research on how algorithmic advice can be presented in order to maximise appreciation towards algorithms and decision aids. It is already shown, that algorithm appreciation is greater when people have the ability to modify its output, even when they see it err (Dietvorst et al., 2016).

Furthermore, understanding numbers and statistics in a technological advanced world is an essential skill. As described by Gigerenzer (2015) a societal problem called collective statistical illiteracy, is not only limited to lay people and patients with low educational background, but also a trend in politics and the medical world in general. To combat this, Gigerenzer and Edwards (2003) suggests presenting natural frequencies instead of single event probabilities, conditional probabilities and relative risks. Since this is shown to improve the amount of correct answers for doctors. It is found interesting to study the effects of data presentation on appreciation of algorithms and decision aids.

Hence, it would be relevant to research other ways in which algorithmic advice can be presented and how decision aids are designed to optimise appreciation towards algorithms and decision aids in the contexts of use. This leads to the following initial research question regarding algorithmic advice presentation:

How are decision aids designed to optimise trust in the given advice?

This chapter has focused the attention on introducing studies regarding trust in algorithmic advice in general, focusing on what is known about algorithm appreciation, the effects of expertise and advice presentation. The forthcoming chapter discusses the research questions and the methods that can be applied during this thesis to answer them and better understand trust in decision aids and algorithms.

3

Research questions

This chapter contains research questions regarding expert professionals' trust in decision aids based on the initialising research questions. As touched upon in the previous chapters of this thesis, it is problematic that a distrust in algorithms, algorithm aversion, might result in decision aids not being optimally used, since algorithms in the right contexts improve judgement accuracy. Which in turn might lower patient security in hospitals, lower road safety, and halt artificial intelligence progress in general. The focus is on discussing the research questions, the methods applied throughout, and the main focus of this thesis.

Main focus area

The healthcare domain is chosen as the focus area due to the large focus on and upcoming implementation of nationwide machine learning based decision aids in the Danish healthcare domain ([Sundheds- og Aeldreministeriet, 2018a](#)). The healthcare domain is currently undergoing a rapid digitisation through sensor technology and machine learning algorithms. This is the case with patient monitoring devices, devices monitoring patient flow, tool management, etc. Furthermore, there is an expectation of a clearly hierarchical structure in the healthcare sector that defines different roles clearly in the healthcare professionals' and patients' views. Hence the main research question follows as:

Do healthcare professionals distrust decision aids and algorithmic advice, and why?

Overall method considerations

To answer the research question it is relevant to understand to what extent expert professionals trust advice from each other, why they might trust each other and what affects this trust. This gives the following research question:

Do healthcare professionals trust advice from other healthcare professionals, and why?

Lastly, in order to answer whether healthcare professionals trust or distrust decision aids and algorithms it is relevant to understand, what can be altered to affect expert professionals' trust in decision aids. Furthermore, it is relevant to explore how these can be designed in order to optimise trust in the advice given by the decision aids. This results in the following research question:

How are decision aids designed to optimise trust in the given advice?

To answer the research questions a quantitative study with the purpose of rating subjective aspects of human expert professionals and decision aids is performed. This results in a mapping of the factors of trust in advice which is analysed through statistical analysis such as probabilistic choice models (BTL, preference trees) or principal component analysis. This technique has previously been used in regard to mapping dimensions in mind perception ([Gray, Gray, & Wegner, 2007](#)). Hence, it is chosen to use a similar method in order to understand expert professionals' perception of decision aids and other expert professionals, regarding trust in advice. This study leads to a comparison of human experts and decision aids, and which factors affect trust and to what extent. Another way to map trust in advice from different advisers is through controlled experiments with the purpose of measuring weight of advice. Weight of advice is measured through the use of a forecasting experiment similar to those seen in the advice taking and decision making literature ([Bonaccio & Dalal, 2006; Logg et al., 2019; Önkäl et al., 2009](#)). These results can also be mapped using a principal component analysis.

To carry out an experimental study as described above, it is necessary to perform an exploratory study which explores the important aspects of trust in decision aids amongst expert

professionals. This can be done through a qualitative study such as an interview study focusing on expert professionals and their opinions and use of decision aids.

To summarise, this study seeks to research whether expert professionals in the Danish healthcare domain are averse to decision aids and algorithmic advice, and how appreciation of advice from decision aids and algorithms are optimised through design and presentation. This is done with the focus of mapping algorithm perception through an exploratory interview study. And through an experimental study comparing different decision aids and human experts based on general statements about trust in decision aids found during the exploratory study.

The following chapter describes how the exploratory interview study is conducted to answer the three research questions.

4

Exploratory study

This chapter contains a description of the exploratory study, the methods used and the analysis of the collected data.

The exploratory study seeks to understand people's perception of decision aids, algorithms and other experts as well as prepare for an experimental study, which compares people's perception of and trust in advice given from different decision aids and other experts. Hence, the purpose of the exploratory study is to gain a general knowledge about people's perception of decision aids, algorithms and other experts. To explore how the perception of decision aids affects appreciation towards them and how healthcare professionals' perception of algorithms differ from their perception of other experts, the focus is on the Danish healthcare domain.

4.1 Preparation for the experimental study

To conduct an experimental study with the purpose of mapping the dimensions of people's perception of decision aids and other experts, it is necessary to explore relevant adviser characters and perceptual statements about these in the healthcare domain. The adviser are patients, other healthcare professionals, administration, IT-professionals, and different decision aid systems. Using characters and descriptions thereof has been used in mapping the dimension of mind (Gray et al., 2007). These characters include Charlie (a Springer spaniel), Sharon Harvey (38 years old and works as an advertiser), and Kismet (a sociable robot). Hence, the physical attributes could consist of different types of advisers such as algorithms, experts and laypeople. The perceptual statements of this study could therefore include a range of perceptual statements with the purpose of mapping healthcare professionals' trust in different adviser characters. This approach is also used in mapping the dimensions of mind (Gray et al., 2007). To explore these adviser characters and perceptual statements about them, an exploratory study is conducted as described in the following sections.

4.2 Considerations about methods

To decide upon the advisers and perceptual statements to be used in the experimental study it is important to have a thorough knowledge about context of the focus area, the Danish healthcare sector. The focus of this study is to find the adviser characters and perceptual statements about these. This is done by understanding the focus area, by focusing on qualitative interviews rather than surveys. Surveys are advantageous in gathering demographic data, or tests which can gather information about individuals' performances (Brinkmann & Tanggaard, 2010). Hence, this study focuses on qualitative approaches to gathering knowledge about the focus area.

A common approach to gathering qualitative information is the interview (Brinkmann & Tanggaard, 2010). Interviews are often used to get access to people's experiences of different phenomena in their world. Hence, interviews make it possible to concentrate on how specific individuals understand certain events, situations, or phenomena in their own life.

Research interviews can range from relatively unstructured with few predetermined questions to strictly structured with many leading questions from the interviewer. Relatively unstructured interviews are often seen in anthropological and sociological studies. An example of this is field studies in which informal conversations with participants in the field are often used with observation. The biggest advantages of field studies and relatively unstructured interviews are that it is possible to get very close to the participant's world, and they can be carried out on exactly the time when the participant has something to tell (Brinkmann & Tanggaard, 2010). This gives field studies a high interpretation and ecological validity, but a low external validity. The risks of unstructured interviews are that they can be difficult to carry out and that they require great training and interpersonal sensitivity from the interviewer. Furthermore, it can be difficult to keep focus on predetermined research questions and interviews can end up pointing in many directions.

The strictly structured interview is known as almost questionnaire-like interviews, as seen in phone interview surveys, in which it can be difficult to nuance answers outside of the pre-determined questions (Brinkmann & Tanggaard, 2010). This includes questions such as 'rate on a scale from 1-5' or categorical answers such as 'very satisfied to very dissatisfied'. Strictly structured interviews are similar to experimental studies in which a low external and ecological validity is often achieved, but a strong internal validity is achieved. These interviews are useful when answers from many respondents are needed but is not necessary during this exploratory study.

In between the relatively unstructured and the strictly structured interview lies the semi-structured interview. A semi-structured interview functions as an interaction between the interviewer and the interview person (Brinkmann & Tanggaard, 2010). This way the interaction is based on the interviewers questions, from which some questions are predetermined, and the answers of the interview person. The semi-structured interview achieves a high external validity but lacks ecological validity. The semi-structured interview will be used in the beginning of this exploratory study in order to gather information about the focus area, Danish healthcare sector, as well as to possibly gather information that can be the basis of descriptions of adviser characters and perceptual statements about these, that can be used for an experimental study.

Another approach to interviewing is *Contextual Inquiry* as described by Holtzblatt and Beyer (2016). Contextual Inquiry consists of in-context interviews, often performed with users, of systems, in this case decision aids. These contextual interviews are a combination of in-the-field observation, unstructured interviewing, and co-interpretation with the interview person along the interview. Hence, the contextual interviews seek to obtain a high amount of external validity through interview while at the same time obtaining a high amount of interpretation and ecological validity from in-the-field observation and co-interpretation with the interview participants. This interview method is best applied when interviewing actual users in their own context. The contextual inquiry method is used during this exploratory study to explore the contexts of healthcare professionals working with decision aids.

4.3 Exploring people and decision aids in the healthcare domain

To explore the people and decision aids used in the Danish healthcare domain it is chosen to conduct three semi-structured interviews. Furthermore, it is chosen to conduct a contextual interview with a clinician working in the North Denmark Region healthcare. The interview designs, including detailed methods, interview subjects, contexts, and procedures, for the semi-structured and contextual interviews are explained in the following subsections.

4.3.1 Semi-structured interview method description

The semi-structured interviews are based on an interview guide. This interview guide consists of research questions and interview questions, similar to that described by Brinkmann and Tanggaard (2010), see section A.1, section B.1, or section C.1 for examples of the interview guides. The research questions are deeper more abstract questions that lead the overall interview direction, whereas the interview questions are bite-sized questions that can be used in order to answer the more abstract research questions. The research questions are not meant to be used directly in the interview but are useful for the interviewer to make sure that they have answered the central points during the interview. Since it is a semi-structured interview, the interview questions are not used word for word during the actual interview but can be used if necessary.

The interview data, the recorded audio file, is transcribed, see section A.2 for an example, and analysed. The transcription is made without an emphasis on tone of voice, intonation, and voice volume, but includes timestamps, pauses, and interjections such as 'ohh' and 'oh-hhm'. The interview is transcribed using a simple transcription strategy, as mentioned, and with a focus on maintaining the meaning of the content. The transcription is carried out using oTranscribe (Bentley, 2019).

The transcribed data is analysed by coding for potential physical and perceptual attributes. This is done by meaning condensation, a systematic text condensation method, on the basis of the research questions from the interview guide. The coding and meaning condensation is made by reading through the transcribed interview using the research questions as guidance for the meaning condensation and coding. To do this, a document is made containing three columns; one for the relevant text transcript section, one for the meaning condensation, and one for categorisation, analysis and interpretation. An example of this is seen in section A.3. Additionally, theoretical reading is used along with the meaning condensation in order to triangulate the

information gathered through the interview. Furthermore, relevant literature suggested by the interview participant is reviewed along the process as well.

4.3.2 Interview I – people in healthcare

The first semi-structured interview has the main purpose of exploring people in the healthcare domain. The interview guide, the transcription, and the meaning condensation of interview I are found in [Appendix A](#).

Interview participant

The semi-structured interview is conducted as an expert interview. The interview participant is Pernille Scholdan Bertelsen, associate professor at Aalborg University, Department of Planning, and specialised in the cultures of the Danish healthcare domain, healthcare informatics and participatory design in the healthcare domain.

It is found that the interview person has a deep knowledge in the domain of healthcare. The interview person has spent the last 20 years doing research in healthcare informatics and information technology in the Danish healthcare sector. The research has primarily focused on structures and work activities in the healthcare domain as well as user involvement in the design of new healthcare technologies or work activities. The interview person has themselves not been working in the healthcare sector, e.g. in hospitals, clinics, etc. The interview person has, however, partaken in field, interview and observation studies in hospitals and clinics. Based on these studies the interview has a solid foundation for answering questions about people in the healthcare domain, their responsibilities, and what characteristics different roles might have. Furthermore, the interview person, has a deep knowledge in society's perception of healthcare informatics, since the interview person has done, field, observational, interview, and questionnaire studies regarding this.

Setup and environment

The semi-structured interview is conducted in a meeting room at the Aalborg University CREATe building at Campus Aalborg, in which the Department of Planning is situated. The interview is conducted from 8.30 am to 9.30 am. Hence, the interview is 1 hour in total, including walk from the interview participant's office to the meeting room as well as introduction and walk back to the office. The recorded interview took a total of 46 minutes and 39 seconds.

Interview procedure

The interview starts at 8.30 am at the office of the interview participant. From here a couple of minutes walk is made towards a small meeting room in which the interview is conducted. Once in the meeting room a short introduction about the interview, including length and purpose, is made to the interview participant. The participant is afterwards informed about anonymity, confidentiality and that the interview is recorded for transcription and analysis purposes, and that these are used in regards to the master's thesis. This introduction was followed by a 46 minutes and 39 seconds recorded, semi-structured interview based on the interview guide, see [section A.1](#). The interview was followed by a walk back to the office of the interview participant at around 9.30 am.

4.3.3 Interview II – decision aids in healthcare

Based on the first interview it is possible to determine roles of healthcare professionals in the Danish healthcare domain and it is possible to extract some perceptual statements regarding trust in advice and perceptions that affect trust. However, it is not possible to extract which decision aids are used in the healthcare domain. To do this it is chosen to interview an expert in decision aids in healthcare. The interview guide, the transcription, and the meaning condensation of interview II are found in [Appendix B](#).

Interview participant

The semi-structured interview is conducted as an expert interview. The interview participant is Stephen Edward Rees, professor at Aalborg University, Faculty of Medicine and Department of Health Science and Technology, and specialised decision aids in the healthcare domain, and healthcare informatics. He has participated in a research project with the purpose of developing

decision aids and intelligent software such as INVENT and BEACON ([Mermaid Care, 2019; Rees, 2011](#)).

The person interviewed has been a part of the Center for Model-based Medical Decision Support on Aalborg University since the beginning in 98'. The interview person is researching decision-support systems related to patients with lowered lung capacities or deeply sedated patients and therefore has an extensive knowledge of physiology and mathematical modelling of physiology. This has resulted in an sub-group called Respiratory and Critical Care group, where the interview person continues to research in the field. From this research, a system for monitoring and providing decision support was developed called BEACON. Based on this research, the interview is able to answer questions about what kinds of decision support systems that exist in the medical world today, how these systems work and what the general opinion is towards these systems.

Setup and environment

The semi-structured interview is conducted in the office of the interview participant in an Aalborg University building at Campus Aalborg, in which the Department of Health Science and Technology is situated. The interview is conducted from 9.00 am to 10.00 am. Hence, the interview is 1 hour in total, including introduction to the interview. The recorded interview took a total of 45 minutes and 22 seconds.

Interview procedure

The interview starts at 9.00 am at the office of the interview participant. A short introduction about the interview, including length and purpose, is made to the interview participant. The participant is afterwards informed about anonymity, confidentiality and that the interview is recorded for transcription and analysis purposes, and that these are used in regards to the master's thesis. This introduction was followed by a 45 minutes and 22 seconds recorded, semi-structured interview based on the interview guide, see [section B.1](#).

4.3.4 Interview III – aided medication prescription and access to clinicians

The third semi-structured interview seeks to gain knowledge about the medication prescription decision aid currently under development by the North Denmark Region. Furthermore, it seeks to get access to clinicians that can participate in contextual interview about their use of decision aids. To do this it is chosen to interview a project manager implementing decision aids in North Denmark Region. The interview guide, the transcription, and the meaning condensation of interview III are found in [Appendix C](#).

Interview participant

The semi-structured interview is conducted as an expert interview. The interview participant is Lisbeth Røhr Kristiansen, project manager in Digitalisering og IT in the North Denmark Region and in charge of pilot testing of a new national prescription decision aid rolled out across all Danish regions.

Setup and environment

The semi-structured interview is conducted in the office of the interview participant in their offices in Aalborg. The interview is conducted from 2.00 pm to 3.00 pm. Hence, the interview is 1 hour in total, including introduction to the interview. The recorded interview took a total of 40 minutes and 30 seconds.

Interview procedure

The interview starts at 2.00 pm at the office of the interview participant. A short introduction about the interview, including length and purpose, is made to the interview participant. The participant is afterwards informed about anonymity, confidentiality and that the interview is recorded for transcription and analysis purposes, and that these are used in regards to the master's thesis. This introduction was followed by a 40 minutes and 30 seconds recorded, semi-structured interview based on the interview guide, see [section C.1](#).

4.3.5 Contextual Interview

The contextual inquiry is conducted based on four main principles, context, partnership, interpretation, and focus as described by Holtzblatt and Beyer (2016). The interview is performed in the context of the clinician while the target activities regarding medication prescription are carried out as retrospective accounts. Hence, it is important to gather data as if the retrospective events are ongoing experiences rather than summary data from earlier experiences. Furthermore, there is a focus on gathering concrete data rather than abstract data throughout the interview. To achieve this it is important to be able to ask the right questions throughout the interview. This is achieved by building the correct relationship from the beginning of the interview.

The interview is structured to develop a master/apprentice relationship between the interview participant (master) and the interviewers (apprentices). This enables the interviewers to be more straightforward and 'nosy' while asking questions. It is important to avoid becoming the polite guest or the decision aid expert throughout the interview to build the optimal partnership for the interview.

By building a master/apprentice partnership during the interview, it is possible to ask nosy questions (e.g. why did you do that?) and share interpretations throughout the interview. By sharing interpretations with the clinician during the interview and listening closely for the responses it is possible to interpret, condense and analyse data along the interview. Hence, it is chosen not to record and transcribe the interview for further analysis. It is instead chosen to take thorough notes throughout the interview, gathering interpretations and key points made by the clinician.

During this interview the focus is on how the current procedure in terms of medication prescription. Therefore the interview seeks to explore and understand how medication prescription is currently carried out and the clinicians attitude towards new decision aids, in terms of both functionality of the decision aid and practical daily use of the decision aid. After the contextual interview an interpretation session is held to explore the data gathered. The interpretation session is held immediately after the contextual interview in order for the data to be fresh in the memory of the interviewers. The focus of the interpretation is to describe how decision aids should work and look, to further design mock-ups to use in the experimental study.

The interview and interpretation notes from the contextual inquiry is seen in Appendix D.

Interview participant

The contextual inquiry is conducted with a clinician from Aalborg Universitetshospital. The participant is a consultant at cardiology at Aalborg Sygehus Syd.

Setup and environment

The contextual inquiry is conducted at the office of the participant in Aalborg. Hence, it is conducted near his workstation, on which the prescription software is running on his computer. The inquiry is conducted from 2.15 pm to 3.15 pm. Hence, the inquiry is 1 hour in total, including introduction to the inquiry.

Interview procedure

The contextual inquiry starts at 2.15 pm at the office of the participant. A short introduction about the interview, including length and focus, is made to the interview participant. The participant is afterwards informed about confidentiality, and that the interviewer will be taking notes based on the interview and that these are used in regards to the master's thesis.

The interview then changes the rules to an observation and discussion based format in which the clinician shares, shows and tells about the target activities. There is a focus on retrospective accounts and co-interpreting along the interview.

The interview closes with an overall interpretation of the learnings about the activities and context in which the interview has taken place. The interview concludes by thanking the interview participant.

4.4 Results and analysis

This section contains the results and analysis for the three semi-structured interviews and the contextual inquiry.

4.4.1 People in healthcare (Interview I)

The analysis is carried out based on the transcription of the audio from the interview, see [Section A.2](#).

Roles in the healthcare domain - adviser characters

According to the interview person the Danish healthcare sector has been studied for many years. This resulted in the finding that there are three main cultures in Danish hospitals, the clinicians (doctors), the hospital nurses, and the administration (management). These three cultures are described by [Eriksen and Ulrichsen \(1991\)](#) as the main three pillars of the Danish hospitals. The study consisted of interviews as well as participatory observations carried out in Danish hospitals. These three cultures have since, according to the interview person, evolved into four cultures with the addition of the IT professionals as well. This corresponds to the results found by [Petersen and Bertelsen \(2012\)](#), during a large observation and interview field study in the Danish hospital sector. This results in the four cultures; the clinician pillar, the nursing pillar, the administration (management) pillar, and the IT professional pillar. In the Danish hospitals the clinicians and the nurses are also represented in the administration, the clinicians by a medical director (lægefaglig direktør), and the nurses by a nursing director (sygeplejefaglig direktør). Based on the interview it is found that in the clinician pillar there is a general structure consisting of executive consultants (ledende overlæger), consultants (overlæger), staff specialists (afdelingslæger), senior registrars (1. reservelæge), and registrars (reservelæger). This structure corresponds to that which is described by [Eriksen and Ulrichsen \(1991\)](#). The nursing structure is described as head nurses (oversygeplejersker), charge nurses (afdelingssygeplejersker), and social and health care assistants. This structure is likewise backed by the findings of [Eriksen and Ulrichsen \(1991\)](#). Although these structures paint a general image of the Danish hospital sector, it is, through the interview, found that the structure is very flat compared to other healthcare systems. The two pillars with their corresponding structures can be seen below:

- Clinician pillar | Lægesøjlen
 - Medical director | Lægefaglig direktør (part of the administration)
 - Executive consultant | Ledende overlæge
 - Consultant | Overlæge
 - Staff specialist | Afdelingslæge
 - Senior registrar | 1. reservelæge
 - Registrar | Reservelæge
- Nursing pillar | Sygeplejerskesøjlen (part of the administration)
 - Nursing director | Sygeplejefaglig direktør
 - Head nurse | Oversygeplejerske
 - Charge nurse | Afdelingssygeplejerske
 - Social and health assistants | Assisterende sygeplejerske/SOSU-assistent

The interview person highlighted that the structure might have an effect on the weighting of advice from the different roles in the hospital. It is mentioned that the head nurses and consultants might, for instance, have an effect on the way patients weigh the advice given by these.

During the interview, medical secretaries (lægesekretærer), medical laboratory assistants, hospital porters, (registered) dietitians, physiotherapists, and chiropractors were mentioned. The medical secretaries have a key role in the Danish hospital sector in which they help the clinicians throughout their work. Furthermore, they function as quality reviewers of the work of the clinician as well as control part of the administration of their work. Some of these roles, the porters and the laboratory assistants, have little to no patient contact and very defined work tasks. Others, dietitians, physiotherapists, and chiropractors have more contact to patients but little contact to clinicians, and nurses. Since these roles might have more specialised knowledge in some fields than clinicians and nurses, it is found interesting to look at how patients might weigh advice given from these roles compared to clinicians and nurses. Dietitians, physiotherapists, and chiropractors have little authority in the actual decision making in the hospitals compared to clinicians and nurses.

The roles described above are considered when descriptions of different adviser characters are to be made.

Characteristics of healthcare professionals

During the interview a focus was to highlight different characteristics of healthcare professionals in different positions. This showed that clinicians have a positivist view on the treatment of patients. Clinicians see themselves and their work grounded in scientific evidence. According to Eriksen and Ulrichsen (1991) clinicians are in a technological reality in which the most important aspect is scientific testing performed with a narrow scientific methodology. In contrast the nursing pillar sees themselves and their work as based in hermeneutics. Nurses work from a holistic viewpoint based on the whole person. They take care of and nurse the patient when they are undergoing and when they have undergone a treatment. Eriksen and Ulrichsen (1991) describes nurses as working in two different, and at times conflicting, realities; a rational and scientifically provable science based testing methodology and an intuition-based tendency to involve factors that are not scientifically provable.

People's perception of healthcare professionals and their advice

The responsibility and liability of decisions and advice given lies on the clinicians (doctors). The clinician is the specialists who has the responsibility of deciding on treatments and medications while they can delegate smaller decisions to the nurses. According to the interview person this could affect the weighting of advice. Another point made was that even though clinicians and nurses function as a team, the patient most of all sees and talks to the nurses. The contact between clinicians and patients are most often limited to ward rounds. According to the interview person this could also affect the weighting of given advice. Especially from the point of view of the patients but also between healthcare professionals. From the patients perspective the clinician is seen as knowing about treatment and symptoms of diseases, but the nurses are seen as more equal and better at communicating with the patients about the treatment and their health conditions.

According to the interview person patients increasingly use the internet for researching their own symptoms, which possibly affects the weight put on advice from healthcare professionals. Often patients use search engines such as google to look up symptoms for their problems, which in turn involves algorithms in the search. These algorithms provide the patient with search results based on their inputted search criteria.

Based on the interview it is found that 78 % of respondents in a representative questionnaire study in Denmark supports the use of IT in the healthcare sector while 8 % is against an increased amount of IT (Tornbjerg & Bertelsen, 2014). Among the most important factors for the supporters of healthcare IT are the increased ability to find relevant information themselves, being able to follow their own course of treatment, not having to physically meet up with the practitioner, that multiple practitioners are able to view files which combats loss of information and results in better treatment, that practitioners are able to view files and information, increased time for physical consultation since e-consultations are faster for smaller things, and easier contact to practitioners. People against IT in healthcare answers that they are worried that physical consultations and contact is replaced by e.g. e-consultations, that consultations are less thorough when not physical, that it is not necessary, and that they worry for safety of data and medication errors. With the physical consultations being the most important factor.

According to the interview person, the most important aspect for the patients and their weight of advice is that they understand the information that they are given. To measure the ability for patients to understand the information they are given, literacy, the ability to read and write, health literacy, the ability to understand own condition and treatments, and e-health literacy, the ability to use and understand healthcare IT (Norman & Skinner, 2006). Furthermore, the interview person mentions that healthcare IT is not developed for the patients that need it the most, the less health and e-health literate people, but developed with the most self-engaged people that would normally do well without the aids. According to Showell and Turner (2013) healthcare technology is most often design for 'people like us', designers, programmers, and IT professionals, while people who are disempowered, disengaged, and disconnected are disadvantaged since they are not included in the design of new technology for various reasons. This is also highlighted in a Danish context through a questionnaire study in which it is found that the people who need it the most are less often included in the development process (Petersen & Bertelsen, 2017).

Decision aids in the Danish healthcare domain

According to the interview person new decision aid systems are underway in the Danish healthcare domain. These new decision aids focus on digital aids for prescription of medicine to

avoid cross-effects and allergies in patients, prediction and learning, and tracing of the elderly through systematic, digital registration of their health data ([Sundheds- og Ældreministeriet, 2018b](#)). The interview person is not familiar with other decision aid systems. Hence, this is a focus of further research.

Healthcare professionals' perception of algorithms and AI

The interview person does not have much knowledge about healthcare professionals view on and perception of algorithms and decision aids. It is suggested to do further research in order to analyse this. This is, according to the interview person best done by talking to the actual healthcare professionals. Hence, it is chosen to conduct interviews with healthcare professionals about algorithms and decision aids.

Initial adviser characters for the experimental study

The initial characters for the experimental study are based on the roles that have been found in the Danish healthcare domain. It is chosen to focus on the clinician pillar and the nursing pillar as well as the new decision aid systems which gives the following initial characters:

- Consultant
- Charge nurse
- Decision aid for prescription of medicine
- Decision aid for prediction of disease prevention and treatment of the elderly

Initially it is chosen to focus on a consultant (overlæge) and a charge nurse (afdelingssygeplejerske) to represent the two pillars in the Danish healthcare domain. It is chosen to begin with two different adviser characters to limit the amount of trials in the experimental study, and to make room for different algorithmic advisers.

Initial perceptual statements for the experimental study

Based on the interview it is possible to extract initial perceptual statements about trust in advice and perceptions that affect trust in different characters. The following perceptual statements have been extracted from the first interview, these are seen in [Table 4.1](#).

Table 4.1: Perceptual statements from interview 1. The perceptual statements are preceded by the characters, e.g. "The character is able to...". For this table the statement are shortened for simplicity.

is able to take intentional actions	kan tage intentionelle beslutninger.
is scientifically grounded.	er videnskabeligt funderet.
actions are based in scientific evidence.	beslutninger er baseret på videnskabelig evidens.
actions are based on intuition.	beslutninger er baseret på intuition.
is able to form a holistic perspective of others.	kan danne et hollistisk perspektiv af andre.
is able to take action based on factors that are not scientifically provable.	kan tage beslutninger baseret på faktorer der ikke er videnskabeligt testbare.
can take action based on multiple parameters.	kan tage beslutninger ud fra adskillige parametre?
can take responsibility for their actions.	kan tage ansvar for sine beslutninger.
has responsibility for their actions.	har ansvar for sine beslutninger.
can explain their reasoning in a way that I will most likely understand.	kan forklare sit ræsonnement på en måde så jeg med højest sandsynlighed vil kunne forstå det.
can explain their reasoning in a way that a layperson will most likely understand.	kan forklare sit ræsonnement på en måde så en lægperson med højest sandsynlighed vil kunne forstå det.

To further understand the different kinds of algorithms in the healthcare sector it was chosen to explore this subject further. This was done through another interview with a professor with the research area of decision aids and development of software implemented in hospitals in the EU. This interview was carried out before planning the interviews with healthcare professionals to have a better understanding of decision aids in the healthcare domain.

4.4.2 Decision aids in healthcare (Interview II)

The analysis is carried out based on the transcription of the audio from the interview, see section B.2.

Algorithms in the healthcare domain

As mentioned, the interview person has been involved in the development of BEACON. BEACON is a open-loop decision support system, that advises the clinician about the patients respiratory status, and what steps should be taken to improve a patients health. This is just one of the systems that currently is used in the medical world. According to the interview person, different systems that work using algorithms are currently used in the medical world. SmartCare and INTELLiVENT are systems used in respiratory medical work, but unlike BEACON, these are closed-loop systems that are based on the same principles of mathematical modelling of physiology, but with control over the respiratory settings. These systems are according to the interview person more regulation systems than decision support systems. Systems like TREAT ([Andreassen, Zalounina Falborg, & Nielsen, 2019](#)) and Glucosafe ([Department of Health Science and Technology, 2019](#)) provides advice about antibiotic treatment of serious infections and insulin therapy respectively.

Open-loop and expertise

According to the interview person a big debate in the development of decision aids, is how transparent the system needs to be. A sub topic of this debate is the fact that in the future, the medical world will lose more expertise and there will be a lack of experienced clinicians (Anæstesiologer). These considerations were used in the development of BEACON. While helping clinicians treat patients, the system is designed to promote expertise in the user, and enable them to learn from the simulations created by the system. The system was designed to have different levels of information. These levels are targeted at different types of medical personnel, and where some needs to have a full in depth look at the patients status, others might only need the basic information to decide whether they should follow the advice given from the system. According to the interview person this provides a certain trust in the systems functionality and usage, as the medical professionals can get just as much information as they like, at any point in the use of the system. In every discussion on algorithms or technology in the medical world, the topic of responsibility comes up. The BEACON system maintains that the doctor or medical professional in charge does still have full responsibility of the patients and their health. However, a function included in the system, allows the person in charge to set up boundaries so that the system only gives advise inside a certain scope. Hereby, other medical personnel can change the settings for the patient, up until a certain limit, where the person in charge then again can evaluate the patients status.

Attitudes towards technology

The interview person defined three types/levels of attitude towards technology in the medical world. The first type is based on a persons attitude towards technology in general. According to the interview person these discussions of whether or not a person likes technology should be changed or directed towards whether or not a patient could benefit from the use of technology. Medical professionals should only be concerned with their patients, and not concerned with technology, but if some kind of technology has the power to help, they should show interest in these. The second type of attitude is based on whether a person thinks that a certain systems functionality is actually able to help a patient. These are very context dependent, but relates back to transparency. The medical professionals need to be sure of what parameters their patient is rated on, and understand as much as possible, how the system gives an answer. The third type is based on a more practical approach where the system needs to fit into the daily practice of medical professionals. If a system disturbs the daily routine, or requires a lot of extra effort to work with, it will have a negative effect on the attitude towards the system. According to the interview person one of the factors that provide a positive effect on attitude is scientific evidence. Medical professionals and administrators alike, need scientifically proven reductions of illnesses or costs for creating a support in new medical technology.

Additional adviser characters for the experimental study

From the interview the decision aids can be categorised into two. Open-loop decision aids with the focus on having an advisory role that enhances the expertise of the clinician or nurse, and

a closed-loop decision aid with the focus on having a regulatory role.

- Open-loop decision aids (rådgivende system / ekspertisefremmende system)
 - Beacon
- Closed-loop decision aids (reguleringssystem)
 - INTELLiVENT
 - SmartCare

Additional perceptual statements for the experimental study

Based on the interview it is possible to extract additional perceptual statements about what affects trust in different types of adviser characters. The following perceptual statements have been extracted from the second interview, these are seen in **Table 4.2**.

Table 4.2: Perceptual statements from interview 2. The perceptual statements are preceded by the characters, e.g. "The character can inform about...". For this table the statement are shortened for simplicity.

can inform about origin of data/information that forms the basis of their decision/advice.	kan tilkendegive oprindelsen af data/informationer der danner grundlag for beslutninger/råd.
can inform about the processes that forms the basis of their decision/advice.	kan tilkendegive processen der danner grundlag for beslutninger/råd.
can enhance its own expertise	kan fremme egen ekspertise.
can enhance expertise in others.	kan fremme ekspertise hos andre.
shows interest in its patients.	viser interesse i dens patienter.
is able to keep information confidential.	kan holde informationer fortrolige.
is scientifically recognized.	er videnskabeligt anerkendt.
can adapt to new groundbreaking knowledge quickly.	kan omstille sig til ny banebrydende viden hurtigt.
can read the room/understands the context it is in.	forstår hvad er sker i rummet/konteksten.
can take critical decisions after critical changes in a patient.	kan tage kritiske beslutninger efter kritiske ændringer i patienter.
understands critical from non critical situations.	kan skelne mellem kritiske og ikke kritiske situationer.
has an economical incentive in decisions made for patients.	har et økonomisk incitament i beslutninger taget for patienter.

4.4.3 Combined results (interview I and II)

This subsection contains an analysis and discussion of the results gathered through the two first interviews. The purpose is to have a combined list of adviser characters and perceptual statements in order to prepare for the experimental study.

Adviser characters for the experimental study

During the first interview the focus was on finding adviser roles of people in the Danish Healthcare domain, this resulted in a definition of the healthcare domain based on four main pillars of which two was found relevant for this study, the clinician pillar and the nursing pillar. It is chosen to represent these pillars by a consultant and a charge nurse. Furthermore, it was found that patients to a larger and larger extent also tends to take an adviser role based on the information they have gathered from internet searches and similar. Additional characters that are able to take advisory roles are briefly mentioned, such as dietitians, physiotherapists, and chiropractors although these have low decision making authority in the hospitals.

The second interview reinforced the findings of the structure in the Danish hospitals. Although it did however add additional nuances to the different roles in the clinician pillar in the context of highly sedated patients.

This results in the following human adviser characters in Danish hospitals:

- Consultant

- Charge nurse
- Patient with own research
- Dietitian, physiotherapist, or chiropractor (low decision making authority)

During the first interview very little focus was on actual decision aids and algorithms in the healthcare domain. It did however result in a discussion about two new decision aids that are being developed for all Danish hospitals as well as general practitioners. As mentioned earlier these new decision aids focus on digital aids for prescription of medicine to avoid, over-dose, cross-effects and allergies in patients and prediction, learning, and tracing of the elderly through systematic, digital registration of their health data ([Sundheds- og Ældreministeriet, 2018b](#)).

The second interview brought new light on decision aids in the healthcare domain with a much more specialised focus. The focus was on decision aids such as BEACON caresystems and INTELLiVENT which are machine learning based decision or regulation aids for patients in ventilators. The two types of systems differ by BEACON being an open-loop system with the focus on aiding decision-making for clinicians and enhancing expertise in the process. This means, however, that open-loop systems require human input and are thus prone to human error. Closed-loop systems such as INTELLiVENT are regulating systems that function without input from the clinicians, and thus does not enhance expertise. According to the interview person healthcare professionals are in general less likely to accept closed-loop systems into their work. Hence, it is found relevant to compare this experimentally through the experimental study.

Due to systems such as INTELLiVENT and BEACON being very specialised decision aids in a given field, it is chosen to focus on the new medicine prescription decision aid being developed by North Denmark Region. This is done to make sure that it is possible to recruit subjects for the experimental study. This gives the following non-human adviser characters for the experimental study:

- Graphical, open-loop prescription medicine decision aid
- Non-graphical, open-loop prescription medicine decision aid
- Graphical, closed-loop prescription medicine decision aid
- Non-graphical, closed-loop prescription medicine decision aid

Hence, the combined list of adviser characters is as follows:

- Consultant
- Charge nurse
- Patient with own research
- Dietitian, physiotherapist, or chiropractor (low decision making authority)
- Graphical, open-loop prescription medicine decision aid
- Non-graphical, open-loop prescription medicine decision aid
- Graphical, closed-loop prescription medicine decision aid
- Non-graphical, closed-loop prescription medicine decision aid

Perceptual statements for the experimental study

From both interviews a set of perceptual statements were extracted, as seen in [Table 4.1](#) and [Table 4.2](#). These are based on the analysis of the interviews, and describe some of the statements that exist regarding clinicians and decision aid systems. However, some of these statements could possibly be describing the same factor, and therefore further analysis is done by comparing the statements from the two interviews. The statements that are similar and/or describe the same underlying factor are categorised in the following [Table 4.3](#). The statements that were categorised with other similar statements were either reformulated into a fitting statement, ie. "*The character is able to form a holistic perspective of other*" and "*The character can take action based on multiple parameters*" were reformulated into "*Gives advice based on an overall picture of the patient*", or the most fitting statement was picked and the others were discarded. The decision to reformulate or discard statements were based on the danish version of the statements as these are the ones expected to be used in the following experimental study.

Table 4.3: A list of the combined statements from the two interviews. The statement are sorted according to relevance to each other. The perceptual statements are preceded by the characters, e.g. "The character is able...". For this table the statement are shortened for simplicity.

English	Danish
is able to take intentional actions	kan tage intentionelle beslutninger.
is scientifically grounded	er videnskabeligt funderet.
-	-
actions are based in scientific evidence	beslutninger er baseret på videnskabelig evidens.
-	-
is scientifically recognized	er videnskabeligt anerkendt
actions are based on intuition	beslutninger er baseret på intuition
-	-
is able to take action based on factors that are not scientifically provable.	kan tage beslutninger baseret på faktorer der ikke er videnskabeligt testbare.
is able to form a holistic perspective of others	kan danne et holistisk perspektiv af andre.
-	-
can take action based on multiple parameters.	kan tage beslutninger ud fra adskillige parametre?
can take responsibility for their actions.	kan tage ansvar for sine beslutninger.
-	-
has responsibility for their actions.	har ansvar for sine beslutninger.
can adapt to new groundbreaking knowledge quickly.	kan omstille sig til ny banebrydende viden hurtigt.
-	-
can enhance its own expertise.	kan fremme egen ekspertise.
shows interest in its patients.	viser interesse i dens patienter.
is able to keep information confidential.	kan holde informationer fortrolige.
can take decisions after critical changes in a patient.	kan tage beslutninger efter kritiske ændringer i patienter.
-	-
understands critical from non critical situations.	kan skelne mellem kritiske og ikke kritiske situationer.
-	-
can read the room/understands the context it is in.	forstår hvad der sker i rummet/konteksten.
can explain their reasoning in a way that I will most likely understand.	kan forklare sit ræsonnement på en måde så jeg med højest sandsynlighed vil kunne forstå det.
-	-
can explain their reasoning in a way that a layperson will most likely understand.	kan forklare sit ræsonnement på en måde så en lægperson med højest sandsynlighed vil kunne forstå det.
-	-
can inform about origin of data/information that forms the basis of their decision/advice.	kan tilkendegive oprindelsen af data/informationer der danner grundlag for beslutninger/råd.
-	-
can inform about the processes that forms the basis of their decision/advice.	kan tilkendegive processen der danner grundlag for beslutninger/råd.
-	-
can enhance expertise in others.	kan fremme ekspertise hos andre.
has an economical incentive in decisions made for patients.	har et økonomisk incitament i beslutninger taget for patienter.

4.4.4 Final statements

Based on the categorisation and reformulation of the perceptual statements, a list of statements were compiled to be used in the following experimental study. The list of statements can be seen [Table 4.4](#)

Table 4.4: A list of the statements combined from the two interviews and reduced to a total of 10. The perceptual statements are preceded by the characters, e.g. "The character gives...". For this table the statement are shortened for simplicity.

English	Danish
gives advice based in scientific evidence	giver råd baseret på videnskabelig evidens.
gives advice based on factors that are not scientifically provable.	giver råd baseret på faktorer der ikke er videnskabeligt testbare.
gives advice based on an overall picture of the patient	giver råd baseret på et samlet billede af patienten
is responsible for their advice	er ansvarlig for sine råd
can explain their reasoning so that it is understandable	kan forklare sit ræsonnement så det er forståeligt
can enhance expertise in others	kan fremme ekspertise hos andre
can enhance own expertise	kan fremme egen ekspertise
has an economical incentive for their advice	har et økonomisk incitament for sine råd
accounts for changes in the needs of the patient	tager højde for ændringer i patientens behov
keeps information confidential	holder informationer fortrolige

4.4.5 Medication prescription decision aid (Interview III)

To further explore the prescription decision aid planned for development during 2019 an interview with a project manager was carried out. This subsection contains the analysis that is carried out based on the transcription of the audio from the interview, see [section C.2](#).

The interview focused on the new prescription decision aid project, and how they have handled designing for healthcare professionals. Different aspects exist in terms of prescription, namely Cave, Interactions, Max dose and High risk medication. All these needs to be accounted for, and one of the biggest problems in designing for healthcare professionals is getting the data into the system. According to the interview person the healthcare professionals in the North Denmark region uses an ongoing diagnose where the actual data is written into the system when the patient is released from the hospital. Through the process of designing the decision aid the healthcare professionals expressed a wish to get decision aids regarding children and pregnant or breastfeeding. The interview person describes trust in the system as being "alpha omega" when designing for healthcare professionals, and that the system needs to fit the healthcare professionals work process.

The interview provided some insight into how the new prescription decision aid would be designed and what factors have the biggest influence in this specific aspect of the medical world. Additionally the interview provided contact to a healthcare professional that has been a part of the development of the new decision aid. An contextual inquiry was decided upon as method for interviewing the healthcare professional as it is an actual user of the current and future system. The contextual inquiry will be described in the next section.

4.4.6 A clinician's context

A interpretation session is conducted based on the notes taken during the contextual inquiry, see [Appendix D](#). The contextual inquiry provided more contextual information about the current situation of prescribing medicine, and how the new decision aid system, as mentioned in [subsection 4.4.5](#), could help clinicians prevent errors. The current situation consists of a system called Clinical Suite, that connects different sub-systems together, so Clinical Suite works by loading the values needed from different sub-systems depending on which kind of information is needed by the clinician. One of these sub-systems is Columna, which is where prescription of medicine occurs. As described by the clinician, the sub-systems doesn't have any cooperation, so that when prescribing medicine, it's not possible to see the patients different health values that could potentially have an effect on which kind of medicine they can get. If a clinician forgets to check a certain value, and the patient isn't able to use the medicine based on this value, it wouldn't be caught before the patient shows no signs of an improving health status.

One of the suggestions made by the clinician was to define certain cases that the clinicians could select, and the system should then know and present which values were important for the clinician to check for solving this case. Additionally, creating an overview of the patients to see

which ones are the most critical, by displaying the most important values would be beneficial to the clinician. As described in [subsection 4.4.5](#), max dose and possible interactions would also be beneficial for the clinician to be made aware of, if a possible typing error occurs.

One of the main problems described by the clinician is the ability to go with a gut feeling when prescribing medicine, which is a difficult thing for a computer to take into account. Here the clinician is able to view the whole patient and not just the medical values that can be measured. Instead, a decision aid should work like a guide or more a memory aid which help the clinician extract the knowledge that already exists inside their head. Clinicians often discuss patients and their health status, to achieve a better result, and they often occur at the office of the clinicians alongside their computers or at conferences done during the day where staff meets up. These decision aids should fit into the daily routines or fit into the systems that they use now and not create additional systems or work. A decision aid should help them remember and guide them, instead of controlling them, and if possible be designed to mimic the discussion with fellow clinicians.

Scenario

For providing the right setting when doing the experimental study, a scenario is created. The scenario will help put the participants of the experimental study in the right context, when evaluating the different characters and statements. The scenario is based off a hypothetical patient and medical information received through interview 3. The medical information consists of different values and can be seen in [Appendix E](#). The scenario was evaluated in a review session done with a medical student, which can be seen in [Appendix F](#). The final scenario can be seen in [Figure 4.1](#).

Sygehistorie

En 78 årig mand er blevet indlagt på sygehuset, du har lægevagten denne dag og har ansvaret for at tage dig af patienten. Patienten er en kompleks medicinsk patient, med symptomerne hovedpine, træthed og svimmelhed hvilket tyder på forhøjet blodtryk. Patienten er tidligere diagnosticeret med og under behandling for type 2 diabetes, hypertension (forhøjet blodtryk) og KOL (Kronisk Obstruktiv Lungesygdom), og patienten har tidligere haft både AMI (blodprop i hjertet) og apopleksi (blodprop i hjernen).

Patienten har på nuværende tidspunkt en lægemiddelplan der består af følgende ordinationer:

- Amlodipin 10 mg, ½ tbl dgl.
- Atorvastatin 20 mg, 1tbl. Dgl
- Clopidogrel 75 mg, 1tbl. Dgl.
- Cozaar 50 mg, 1tbl. Dgl.
- Gemadol Retard 50 mg, 2 tabletter daglig
- Kaleorid 750 mg, 1 tablet daglig
- Metformin 500 mg, 1 tablet 2 gange daglig
- Ultibro breezhaler 1sug daglig
- Bricanyl turbohaler 1sug ved behov

Figure 4.1: Medical case history used in the experimental study.

Final adviser characters and profiles

As described in [subsection 4.4.3](#), the final list of advisers consisted of eight characters, four healthcare professional characters and four decision aid characters. Based on both the additional interview as well as the contextual inquiry, these were cut down to three characters, the first character being the consultant and the two others, two different decision aids. The choice of cutting the number of characters to three was based on the fact that the additional interview and contextual inquiry provided information regarding these three types of characters and the amount of characters used in the experimental study needed to be reduced.

Therefore the three characters used in the following experimental study will be:

- Consultant
- Non-argumentative prescription medicine decision aid
- Argumentative prescription medicine decision aid

For describing each of the characters, three profiles were created. The inspiration for these profiles originates from the additional interview and contextual inquiry. Names and descriptive

information are entirely made up. These profiles were also put through the review session conducted with a medical student. All three of the profiles included a visualisation of the specific character, descriptive information and a advice given by the character which is linked to the scenario.

The visualisation of the decision aids were graphical interfaces as seen in [Figure 4.2](#) and [Figure 4.3](#). These are based on inspiration drawn from the existing systems used today in the North Denmark healthcare sector, as seen in the contextual inquiry. Both interfaces were designed using Affinity Designer 1.6.5.135.

The screenshot shows a web-based decision aid interface titled 'PrescriptAid - Medicinordinering'. At the top, there are two input fields: 'Præparatets navn:' containing 'Rasilez' and 'Generisk navn:' containing 'Aliskiren'. Below these is a search bar with placeholder text 'Indtast søgeord' and a magnifying glass icon. A table lists four drug preparations: Rasilez (Aliskiren, Tablet), Cozaar (Losartan, Tablet), Ancozan (Losartan, Tablet), and Klomentan (Losartan, Tablet). To the right of the table, a callout box provides advice: 'Undgå Rasilez, forøg eventuel dosis af Cozaar i stedet'. It also includes a section 'Støtte til beslutning:' with information about diabetes, interactions with Metformin, and dosage details for Cozaar (50 mg, 1 tablet daily; 100 mg, 1 tablet daily). At the bottom, there are buttons for 'ANNULLÉR' and 'ORDINER'.

Præparatets navn	Generisk navn	Form
Rasilez	Aliskiren	Tablet
Cozaar	Losartan	Tablet
Ancozan	Losartan	Tablet
Klomentan	Losartan	Tablet

Figure 4.2: Shows the decision aid called PrescriptAid.

The main difference between the two graphical interfaces, is that first one called PrescripAid is the argumentative decision aid whereas the second called PharmaCision is the non-argumentative. This difference is seen in both the advice given, and in the search function. PrescripAid tries to explain the advice based on diagnosis and current medical prescriptions and shows similar drugs to the one searched for through the search engine. PharmaCision shows the advice and restricts the user until the advice is followed, and shows only drugs that match the actual word the user has searched for.

The screenshot shows a search interface for pharmaceutical products. At the top, there are two input fields: 'Præparatets navn:' containing 'Rasilez' and 'Generisk navn:' containing 'Indtast søgeord'. Below these is a table with four rows, each representing a product. The columns are 'Præparatets navn', 'Generisk navn', and 'Form'. The products listed are:

Præparatets navn	Generisk navn	Form
Rasilez	Aliskiren	Tablet
Rasagilin "Accord"	Rasagilin	Tablet
Ramipril "Actavis"	Ramipril	Tablet
Raloxifen "Teva"	Raloxifen	Tablet

Below the table, it says 'Viser 4 af 4 præparer'. A red banner at the bottom contains the text 'Forøg i stedet dosis af Cozaar' with a red 'X' icon on both sides. At the very bottom are two buttons: 'ANNULLÉR' and 'ORDINER'.

Figure 4.3: Shows the decision aid PharmaCision.

This chapter resulted in a list of adviser characters and perceptual statements about these from the healthcare domain. Furthermore, the chapter contains the development of decision aids and descriptions of these and a consultant that are able to be used as stimuli in an experimental study. The purpose of which is measuring how healthcare professionals perceive decision aids and other healthcare professionals, as well as how decision aids can be designed to optimise trust in them based on the perceptual statements. How this is carried out is described in the following chapter.

5

Experimental Study

This chapter contains a description of the experimental study carried out as part of this thesis. Furthermore, this chapter contains the results and data analysis of the experimental study.

5.1 Study description

The main goals of this experimental study are to study; from whom clinicians prefer to take advice, which perceptions affect whether clinicians prefer advice, and how this can be used to optimise the trust in decision aids as described in [chapter 3](#). To study this an experimental study is carried out based on the methods and experimental design described in the following sections.

5.2 Method considerations

In order to answer the main research questions it is chosen to carry out an experimental study. A way to answer the research questions is to carry out a multidimensional scaling (MDS) experiment such as the experiment seen by [Petiot and Grognet \(2002\)](#) or [Gray et al. \(2007\)](#). Multidimensional scaling can be used to visualise a complex set of data. By conducting a MDS each stimulus is represented as a point in an n-dimensional space, with n representing the number of dimensions. The distance between points in a dimension explain the similarity of the stimuli in this dimension. With this approach it is possible to reveal "hidden", perceptual dimensions that are used in subjects' judgement of the stimuli.

[Petiot and Grognet \(2002\)](#) shows how MDS is used to find subjective attributes that affect the preference of car models. [Gray et al. \(2007\)](#) uses principal component analysis (PCA) to find dimensions of people's perception of mind. Multidimensional scaling has been widely used in sound and acoustics as well as measuring food preferences, from loud speaker design ([Choisel & Wickelmaier, 2005](#)) to ambient light colours' effect on wine preferences ([Oberfeld, Hecht, Allendorf, & Wickelmaier, 2009](#)).

To perform multidimensional scaling the input data needed are a similarity matrix, consisting of data explaining the similarity of the stimuli, measurements of subjective attributes for each stimuli, and lastly a preference matrix, consisting of data of the preferred stimuli. Two ways of gathering a similarity matrix are to conduct a pairwise comparison study of the stimuli, as seen by [Petiot and Grognet \(2002\)](#) or to conduct a regular scaling experiment of the stimuli. Afterwards, a scaling experiment can be conducted to rate each stimuli based on subjective attributes. Lastly, a paired comparison, scaling or similar study is able to determine which subjective attributes are preferred by the subjects, users or consumers.

The resulting data from a scaling study is able to be analysed using principal component analysis, an example of how the final analysis could look using principal component analysis can be seen on [Figure 5.1](#). This could be done since it makes it possible to analyse which subjective attributes affect clinicians preferences when receiving advice. This can be carried out as a simple pairwise comparison experiment or as a scaling experiment. Another possible method for measuring which adviser is preferred by clinicians is to conduct a Weight of Advice experiment ([Dietvorst et al., 2015, 2016; Logg et al., 2019; Promberger & Baron, 2006](#)). This way, it is possible to measure to what extent clinicians actually weigh the advice given from different advisers. Hence, this method would seek to remove the subjective assessment present when conducting a simple pairwise comparison study or a scaling experiment. Initially this study is not able to map a preference to the principal component analysis, although it is possible to conduct further study into this. Due to the abstractness of the statements to be rated and the difficulty of rating this on scales such as a visual analogue scale, it is chosen not to perform a scaling experiment and not to analyse the data using PCA.

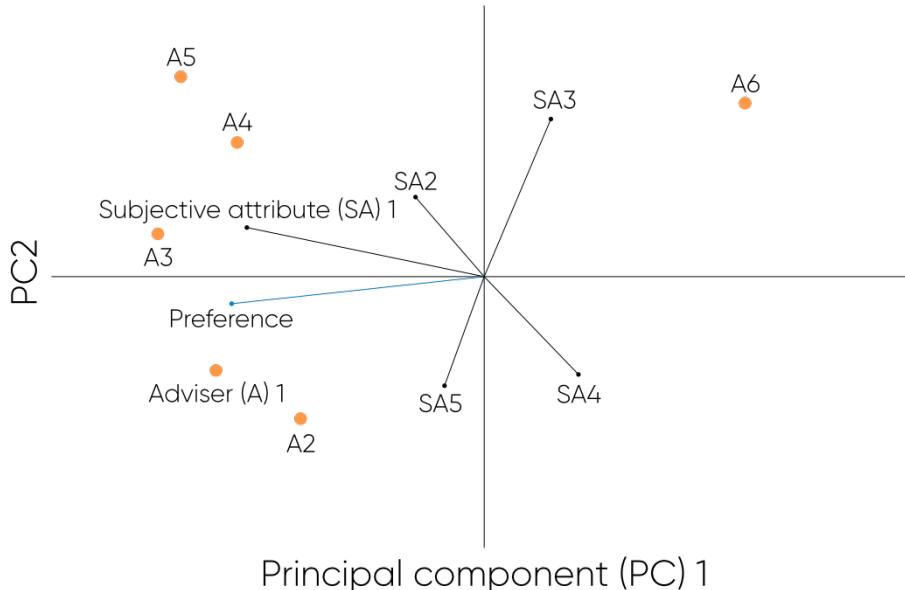


Figure 5.1: Shows an example of how the principal component analysis might explain the advisers (characters), and which subjective attributes (statements) describe the preferred advice

5.2.1 Pairwise comparison assessment and probabilistic choice models

A way to reduce conscious subjective assessment of the attribute ratings and rating of preferred advice is to conduct a pairwise comparison study that forces the subjects to choose between statements or equal advice from different advisers in a pairwise comparison. Paired comparison studies are used to measure people's preferences in regards to sound quality (Choisel & Wickelmaier, 2005; Zimmer & Ellermeier, 2003), unpleasantness of auditory stimuli (Ellermeier, Mader, & Daniel, 2004; Zimmer, Ellermeier, & Schmid, 2004), ambient lights effect on preference of wine (Oberfeld et al., 2009), sounds' ability to prevent right-hand turn accidents for truck drivers (Møller, 2016), and which famous person people would prefer a conversation with (Rumelhart & Greeno, 1971). Common for these studies is that the data are analysed using probabilistic choice models. A probabilistic choice model seeks to model the choices behind answers given by the test subjects (Zimmer & Ellermeier, 2003). Two important requirements of probabilistic choice models are *goodness of fit*, how well the model fits, and *transitivity*, consistency of the answers. If the model has a good fit, it is possible to construct a ratio scale. In order to construct a probabilistic choice model it is necessary to have six test subjects per number of stimuli and have all stimuli compared to each other pairwise.

Two common probabilistic choice models are *BTL* (Bradley-Terry-Luce) model and *EBL* (elimination-by-aspect often used in the form of *Preference Tree* models (Zimmer & Ellermeier, 2003). The BTL model is a simple model structure in which only one attribute distinguishes the stimuli. The EBA (Elimination-by-aspect) model, which is complex, does not necessarily contain a common attribute between all stimuli. The preference tree model is an attempt at a simplified version of the EBA model which contains subgroups of attributes, but with a common attribute for all stimuli. A visualisation of the BTL and preference tree models is seen on Figure 5.2.

Consistency

As previously mentioned consistency of answers is required in order to use probabilistic choice models. Individual data set consistency is analysed by finding the amount of transitivity violations. This means testing for the amount of inconsistent data or the amount of circular triads ($A>B, B>C$, but $A<C$). A conservative estimation for whether the number of transitivity violations is larger than what could be expected by random is by performing a χ^2 test. The individual data sets that are consistent are pooled together into a *cumulative preference matrix*. Furthermore, consistency in the individual data sets does not necessarily result in consistency in the cumulative data. Hence, it is necessary to check for stochastic transitivity in the cumulative data. Stochastic transitivity is checked by checking for weak (WST), moderate (MST) and strong (SST) stochastic transitivity violations. Stochastic transitivity is checked by the following equations:

If Equation 5.1 is true then Equation 5.2, Equation 5.3, and Equation 5.4 checks whether

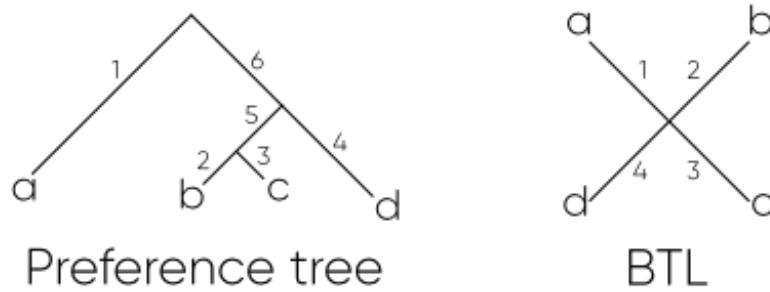


Figure 5.2: Shows an example of the stimuli structure of a preference tree model (left) and a BTL model (right) numbers represent correspond to scale values and letters correspond to stimuli.

WST, MST, and SST hold respectively, where p_{xy} is the probability that x is preferred over y (Krantz, Luce, Suppes, & Tversky, 1971).

$$p_{ab} \geq 0.5 \text{ and } p_{bc} \geq 0.5 \quad (5.1)$$

$$\text{if } p_{ac} \geq 0.5 \text{ WST holds and is violated if } p_{ac} \leq 0.5 \quad (5.2)$$

$$\text{if } p_{ac} \geq \min(p_{ab}; p_{bc}) \text{ MST holds and is violated if } p_{ac} \leq \min(p_{ab}; p_{bc}) \quad (5.3)$$

$$\text{if } p_{ac} \geq \max(p_{ab}; p_{bc}) \text{ SST holds and is violated if } p_{ac} \leq \max(p_{ab}; p_{bc}) \quad (5.4)$$

The consistency in the data determines which model is used to analyse the data. There are no rules for which models can be used, but the general consensus is that none or a low number of SST violations makes it possible to use a BTL model (Tversky, 1972). If there are some SST violations but a low number of MST violations it is possible to use EBA models such as preference trees, and with quite some SST and MST violations and a low number of WST violations it is possible to rank order the data. The amount of stochastic transitivity violations can be compared to the total number of potential triads (Zimmer & Ellermeier, 2003).

Scale value estimation

The ratio scale values of the BTL and EBA models are estimated using maximum likelihood estimation. This means finding the values that give a large probability (likelihood L) for the given answers. The model can be seen in Equation 5.5, where D is the data (frequencies of preference), p_{ij} is the probability of choosing i over j, n_{ij} is the number of observations where stimulus in row i is chosen over stimulus in column j, and n is the number of comparisons for each stimulus pair.

$$L(D|\theta_{model}) = \prod_{i < j} p_{ij}^{n_{ij}} \cdot (1 - p_{ij})^{n - n_{ij}} \quad (5.5)$$

An example of how the ratio scale values of a BTL model might look is seen in Figure 5.3.

Goodness of fit

The second requirement for a probabilistic choice model to hold is goodness of fit. Goodness of fit is estimated by a χ^2 test comparing the results of the model with an unrestricted statistical model that fits the data perfectly. With the assumption of a binomial distribution of each cell value with probability given by the input frequencies. If the model (EBA or BTL) is not significantly worse than the unrestricted statistical model (typically $p > 0.1$), then the model is accepted.

A pairwise comparison study is disadvantageous due to the large amount of subjects needed in order to obtain valid data, six subjects are needed per additional stimuli. Furthermore, a pairwise comparison study is sensitive to answers between subjects not being consistent resulting in violations of stochastic transitivity, especially with a low amount of subjects. It is however chosen to perform a pairwise comparison study instead of a scaling experiment in order to reduce conscious subjective assessment of the statements.

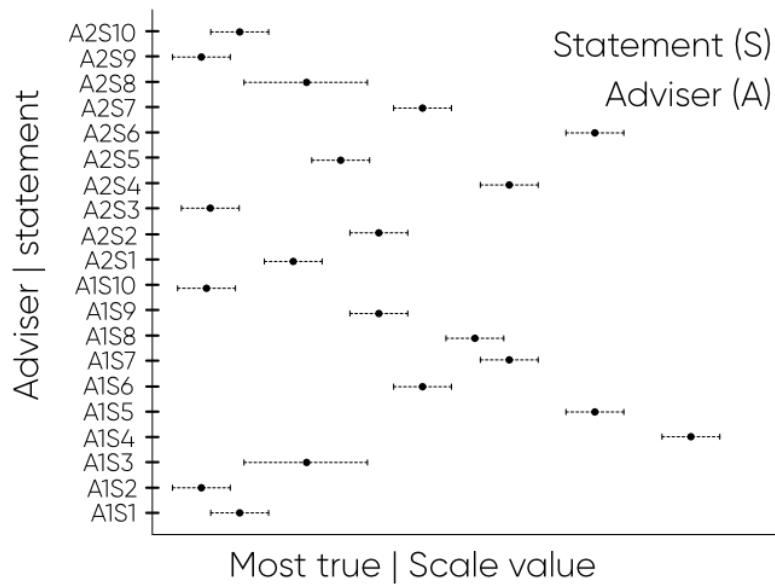


Figure 5.3: Shows an example of how the scale values of a BTL model can be visualised as ratio scale values to understand what is seen as most true by the participants.

To answer the research questions of this thesis, *do healthcare professionals distrust decision aids, and why?*, *do healthcare professionals trust advice from other healthcare professionals, and why?*, and *How are decision aids designed to optimise trust in the given advice?*. It is chosen to perform a pairwise comparison study, in which the participants through paired comparison rate which statements are most true. This way it is possible to study whether there is a difference in healthcare professionals' perception of advice from other healthcare professionals and decision aids. Furthermore, it is possible to gather a subjective attribute matrix of which attributes are chosen the most by the participants.

5.3 Experimental design

As mentioned, this experimental study consists of a pairwise comparison experiment that seeks to measure clinicians' preferences when given advice. Furthermore, the study seeks to answer which subjective attributes might explain these preferences. This is measured through a pairwise comparison study conducted with clinicians, medicine students, other healthcare professional students, and university students in general. The experimental design, including detailed method description, subjects, context, procedure, variables and stimuli for pairwise comparison experiment are explained in the following subsections.

5.3.1 Hypotheses

This subsection contains the suggested hypotheses for the experimental study according to the studies previously reviewed.

According to studies showing a trust towards human advisers, experts especially, compared to decision aids (Önkal et al., 2009; Promberger & Baron, 2006) and anecdotal expressions of the same (Einhorn, 1986; Grove & Meehl, 1996; Highhouse, 2008; Meehl, 1954) the following hypothesis is suggested:

Hypothesis I

Healthcare professionals will choose statements for the consultant to be true to a larger extent than for the decision aid.

According to studies showing a trust towards decision aids compared to human advisers (Dietvorst et al., 2015; Logg et al., 2019; Prahl & Van Swol, 2017) the opposite hypothesis is suggested:

Hypothesis II

Healthcare professionals will choose statements for the decision aid to be true to a larger extent than for consultant.

Lastly, as suggested by Logg et al. (2019) a general distrust towards all advice could be expected from expert professionals, which suggests the following hypothesis:

Hypothesis III

The healthcare professionals will not provide consistent answers, both individually and amongst each other.

How these hypothesis are rejected or confirmed through the experimental design, is explained in the following subsections.

5.3.2 Method description

As mentioned previously it is chosen to conduct pairwise comparison experiment with the purpose of gathering information about the participants' perception of the perceptual statements for each advisor. It is chosen to conduct a pairwise comparison experiment in which the stimuli, the advisers found, developed and described in the exploratory study, and the perceptual statements from the exploratory study are pairwise compared. The statements and characters in the pairwise comparison study are presented as "Consultant Asger Knudsen... ...gives advices based in scientific evidence" and "PrescriptAid... can explain their reasoning so that it is understandable" (Overlæge Asger Knudsen... ...giver råd baseret på videnskabelig evidens - PrescriptAid... ...kan forklare sit ræsonnement så det er forståeligt). The participants are asked to "choose the option that is most true out of the two options" (Vælg den af de to muligheder, der er mest rigtig). An example of a comparison from the experiment can be seen on Figure 5.4.



Figure 5.4: Shows an example of a pairwise comparison from the actual pairwise comparison experiment.

To perform the pairwise comparison study, it is chosen to not have repetitions of the presented statements. Having a repetition would make the experiment take too long time for the participants to be able to concentrate. It is assumed that there is no or very little order effect on whether the statements presented are showed at the top or the bottom. Hence, it is chosen to only show each pair of sentences once. Due to the complexity of the statements transitivity is not assumed, hence it is chosen to compare all combinations of pairs to each other once and not use incomplete blocks.

5.3.3 Final changes based on internal testing and pilot experiment

Due to the final advisor characters being two decision aids and one human advisor, it was decided to remove PharmaCision as an adviser character, and continue with the argumentative decision aid and the consultant. The second decision aid was removed since it was found to be a disadvantage to have an uneven distribution of advisers. In order to have two different decision aids, two human adviser characters would also be needed. This would result in the experimental study taking up too much time and being too exhaustive for the participants to go through. Hence the final adviser characters used in the experimental study can be seen in [Appendix G](#) and [Appendix H](#).

Based on the pilot experiment done with five participants, some different issues were found. The statement '...kan fremme egen ekspertise' (can enhance own expertise) was changed to '...kan omstille sig til ny viden' (can adapt to new information). The rest of the results from the pilot experiment can be seen in [Appendix J](#). Therefore the final statements used for the experimental study can be seen in [Table 5.1](#).

Table 5.1: The final list of statements as used in the experimental study and corrected after pilot experiment feedback. Each statement is represented twice, once per adviser character.

English	Danish
gives advice based in scientific evidence	giver råd baseret på videnskabelig evidens.
gives advice based on factors that are not scientifically provable.	giver råd baseret på faktorer der ikke er videnskabeligt testbare.
gives advice based on an overall picture of the patient	giver råd baseret på et samlet billede af patienten
is responsible for their advice	er ansvarlig for sine råd
can explain their reasoning so that it is understandable	kan forklare sit ræsonnement så det er forståeligt
can enhance expertise in others	kan fremme ekspertise hos andre
can adapt to new knowledge	kan omstille sig til ny viden
has an economical incentive for their advice	har et økonomisk incitament for sine råd
accounts for changes in the needs of the patient	tager højde for ændringer i patientens behov
keeps information confidential	holder informationer fortrolige

5.3.4 Experimental subjects

The pairwise experiment and analysis is conducted with 28 subjects with a mean age of 24.3 ranging from 20–64 years old, 2 clinicians with a mean age of 47 ranging from 30–64 years old, 23 medical students with a mean age of 22.4 ranging from 20–27 years old, two students from medicine with industrial specialisation with an age of 21 years old, and a charge nurse at 27 years old. Additionally the experiment was conducted on 5 students from non-medicine relevant study programmes, these were excluded due to them having difficulty performing the experiment. Hence, these are not represented in the total and mean age of the participants. The mean age of the 5 university students from different degrees other than medicine and excluded from the analysis was 25 ranging from 23–26 years old. It is chosen to conduct the pairwise comparison study with clinicians since this is the main users of the new decision aid systems developed for the Danish regions. Furthermore, it is chosen to use medical students, preferably with practical experience with patients and decision making from field work, student jobs and/or rota (turnus). To find additional participants it is chosen to also include other healthcare professionals as well as university students from other degrees than medicine (later excluded). This is done in order to be able to have enough participants for the pairwise comparison study.

It is chosen to gather demographic data about the participants as part of an exit interview after having performed the pairwise comparisons. These questions are about age, current position, if student the semester, if working the amount of work experience in healthcare, type of experience, job title, and gender.

5.3.5 Setup and environment

The following subsections describe the materials of the pairwise experiment as well as the experimental setup.

Materials

- ASUS zenbook UX 360U
- Application with pairwise comparison experiment, see [Appendix I](#) and electronic appendix [section L.1](#)
- Descriptions of characters (on paper), see [Appendix H](#) and [Appendix G](#)
- Description of a scenario and medical information (on paper), see [subsection 4.4.6](#) and [Appendix E](#)
- Declaration of consent form (on paper)

Experimental setup

The experiment is carried out at the location of the subjects, this means that the experiment is carried out in offices of Aalborg University Hospital, Aalborg University group rooms with medical students and other students, and at the offices of clinicians in clinics around Aalborg.

The computer is set up with the program already running in full screen. The subjects are able to perform the pairwise comparison study on the touch screen of the computer directly in the program.

The data from the experiment is saved locally onto the computer throughout the experiment as described in [Appendix I](#).

Stimuli

The stimuli presented to the participants are the 20 statements, 10 for each character, see [Table 5.1](#) as well as the descriptions of the two adviser characters. The list of statement is revised on the basis of the comments received during the pilot experiment, see [Appendix J](#). The adviser characters are presented to the participants on paper. This gives the participants the possibility to always reference the descriptions while performing the pairwise comparisons. The character descriptions can be seen in [Appendix G](#) and [Appendix H](#). Furthermore, the participants are instructed to identify with being a doctor in a given scenario, see [Figure 4.1](#).

Having a paired comparison of 20 statements without a repetition and without the same statement for the same adviser character being compared with itself gives a total of 190 comparisons. This is illustrated in [Equation 5.6](#), where s is the amount of statements and c is the amount of adviser characters.

$$\frac{((s \cdot c) - 1)(s \cdot c)}{2} = \frac{((10 \cdot 2) - 1)(10 \cdot 2)}{2} = 190 \quad (5.6)$$

5.3.6 Variables

Pairwise comparison variables

The independent variables of the pairwise comparison experiment are the 20 statements made up of 10 for each of the 2 characters as well as the descriptions of the characters. The dependent variable is the choices made in the pairwise comparison of the 20 statements. Hence, the dependent variable is a pairwise comparison of which statements the participant thinks are most true (mest rigtig).

Exit interview variables

In the end of the experiment the participants perform a short exit interview. The main purpose of the exit interview is to gather demographic data about the participants. The independent variables in the questionnaire are the actual questions, which can be seen in [section L.3](#). The dependent variables are the answers to the questions. The purpose of gathering demographic data about the participants is to be able to see whether there are differences in the results that can be explained by the demographic data.

5.3.7 Experimental procedure

The experimental procedure for the pairwise comparison experiment and exit interview is as follows:

1. Introduction to the study according to the script, see [Appendix K](#).

- a) The participant is introduced to the scenario and medical information, see [Figure 4.1](#) and [Appendix E](#).
- b) The participant is introduced to the adviser character descriptions, see [Appendix G](#) and [Appendix H](#).
- 2. The declaration of consent is signed on paper.
- 3. The pairwise comparison experiment begins.
 - a) An on-screen introduction screen is shown.
 - b) The pairwise comparisons are performed.
 - i. Two statements are compared at a time, the character descriptions are always present, and it is not allowed to return to previous comparison.
 - c) After 10 minutes a small break is held in which the participant is offered cake/cookies/fruit and coffee/tea/water.
 - d) The remaining pairwise comparisons are performed.
 - e) When the last comparison is made, the data is saved.
 - f) The participant is shown an end screen.
- 4. The exit interview begins.
 - a) One question is asked at a time according to the question list.
 - b) The answers are noted down in a Google sheets spread sheet.
- 5. Debriefing is performed during the exit interview, see electronic appendix [section L.3](#).

During the pilot test of the experiment the time stamps for the experiment were as follows: Approximately 6–7 minutes for introduction, approximately 21–25 minutes for pairwise comparison, and approximately 2–5 minutes of exit-interview. The total time for the pilot test ranged between 29 minutes and 20 seconds and 35 minutes and 22 seconds.

5.4 Results and data analysis

This section contains the results gathered throughout the experimental study. The exit interview data is seen in the electronic appendix and the raw data from each subject is also found as an electronic appendix, see [section L.2](#), and [section L.3](#). The data analysis is performed in RStudio version 1.0.143 ([RStudio, 2019](#)) using the eba package ([Wickelmaier, 2019](#); [Wickelmaier & Schmid, 2004](#)). The R script used for data analysis can be found in electronic appendix [section L.4](#).

5.4.1 Consistency

Transitivity of individual subjects

The consistency of the individual data sets is checked separately for each individual. This is done by counting the amount of circular triads in each individual data set. The 28 subjects produced a median number of 51 circular triads out of a possible 330 and none of the subjects exceeded the amount expected by chance (X^2 -test $\alpha = 0.05$). Hence, by a conservative estimate all participants gave consistent answers to be included in further analysis.

Stochastic transitivity

The data is pooled into cumulative matrix which is seen in [Table 5.2](#). The cell entries indicate how many subjects judged the statement in the given row as more true (rigtig), than the statement in the respective column. Row and column numbers correspond to the following list of statements:

Table 5.2: Cumulative preference matrix. The cell entries indicate how many subjects judged the statement in the given row as more true (rigtig), than the statement in the respective column. Numbers 1–10 represent statements with the consultant 11–20 represent statements for PrescriptAid, e.g. 1./11. Overlæge Asger Knudsen/PrescriptAid giver råd baseret på videnskabelig evidens. The following are the remaining sentences: 2./12. ...giver råd baseret på faktorer der ikke er videnskabeligt testbare. 3./13. ...giver råd baseret på et samlet billede af patienten. 4./14. ...er ansvarlig for sine råd. 5./15. ...kan forklare sit ræsonnement så det er forståeligt. 6./16. ...kan fremme ekspertise hos andre. 7./17. ...kan omstille sig til ny viden 8./18. ...har et økonomisk incitament for sine råd. 9./19. ...tager højde for ændringer i patientens behov. 10./20. ...holder informationer fortrolige.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	18	3	8	6	11	8	24	2	15	3	26	13	18	16	12	11	19	15	18
2	10	–	4	4	6	7	6	25	7	12	4	26	14	17	14	12	4	15	15	17
3	25	24	–	22	21	26	25	28	22	25	10	27	21	25	25	23	19	26	19	21
4	20	24	6	–	15	18	20	28	11	21	12	28	18	25	22	17	14	24	22	25
5	22	22	7	13	–	17	20	27	5	20	10	27	20	25	23	20	12	25	22	23
6	17	21	2	10	11	–	15	25	8	16	6	25	14	18	18	16	11	19	16	20
7	20	22	3	8	8	13	–	27	4	20	4	28	17	22	21	14	10	19	17	20
8	4	3	0	0	1	3	1	–	0	7	1	14	3	7	2	1	4	5	5	2
9	26	21	6	17	23	20	24	28	–	20	10	28	24	25	23	22	19	24	21	24
10	13	16	3	7	8	12	8	21	8	–	6	25	15	18	15	12	8	16	12	17
11	25	24	18	16	18	22	24	27	18	22	–	27	21	26	26	24	19	25	24	25
12	2	2	1	0	1	3	0	14	0	3	1	–	5	7	0	1	0	7	4	3
13	15	14	7	10	8	14	11	25	4	13	7	23	–	19	15	10	6	15	17	17
14	10	11	3	3	3	10	6	21	3	10	2	21	9	–	9	6	5	16	10	12
15	12	14	3	6	5	10	7	26	5	13	2	28	13	19	–	5	9	18	16	13
16	16	16	5	11	8	12	14	27	6	16	4	27	18	22	23	–	10	19	18	19
17	17	24	9	14	16	17	18	24	9	20	9	28	22	23	19	18	–	20	21	25
18	9	13	2	4	3	9	9	23	4	12	3	21	13	12	10	9	8	–	9	15
19	13	13	9	6	6	12	11	23	7	16	4	24	11	18	12	10	7	19	–	12
20	10	11	7	3	5	8	8	26	4	11	3	25	11	16	15	9	3	13	16	–

It is chosen to perform a transitivity test on the pooled probabilities, as described in section 5.2. The pooled data set is checked for weak, moderate and strong stochastic transitivity violations (WST, MST, SST). The cumulative data matrix shows that out of a possible of 1140 possible, 8 WST, 33 MST, and 281 SST violations are found, which can be seen in Table 5.3.

Table 5.3: Number of stochastic violations and total triads.

	WST	MST	SST	Total triads
Numeric values	8	33	281	1140
Percent value	0.7%	2.9%	24.65%	

Therefore, the amount of WST considered low when comparing the amount to the amount of possible violations, this also holds true when compared to other studies (Choisel & Wickelmaier, 2005; Ellermeier et al., 2004; Møller, 2016). Although many studies have results showing 0 WST violations (Choisel & Wickelmaier, 2005; Møller, 2016; Zimmer et al., 2004), these are out of a lower amount of total possible violations, between 56 and 220. This implies that weak stochastic transitivity holds and that it is possible to establish a uni-dimensional ordering of the stimuli with respect to how true, the statements are found. When compared to other studies the amount of MST violations is also found to be low compared to the total amount of possible violations, as well as compared to other studies (Choisel & Wickelmaier, 2005; Ellermeier et al., 2004; Møller, 2016; Zimmer et al., 2004). This indicates that probabilistic choice models such as the BTL or EBA (preference tree) models may be fit to represent the data. In order for a BTL model to represent the data, a low amount of SST violations are required. 281 SST violations out of 1140 possible suggests that the BTL model may not hold, although less restrictive models such as preference trees might. The amount of SST violations is also seen as high, when compared to other studies (Choisel & Wickelmaier, 2005; Ellermeier et al., 2004; Møller, 2016; Zimmer et al., 2004). There is no statistical tests available to test for the significance of these violations (Zimmer & Ellermeier, 2003). It is important to evaluate the fits of the various probabilistic choice models.

5.4.2 BTL model

Despite what is suspected by the amount of SST violations the likelihood-ratio test for the fit of the BTL model did not indicate significant difference from the model prediction $X^2(171) = 150.8$; $p = 0.864$. It is therefore concluded that the BTL model holds, which indicates that the underlying perceived correctness of the statements is uni-dimensional. The data are therefore represented on a preference ratio scale as normalised utility scale values seen on Figure 5.5 and in Table 5.4.

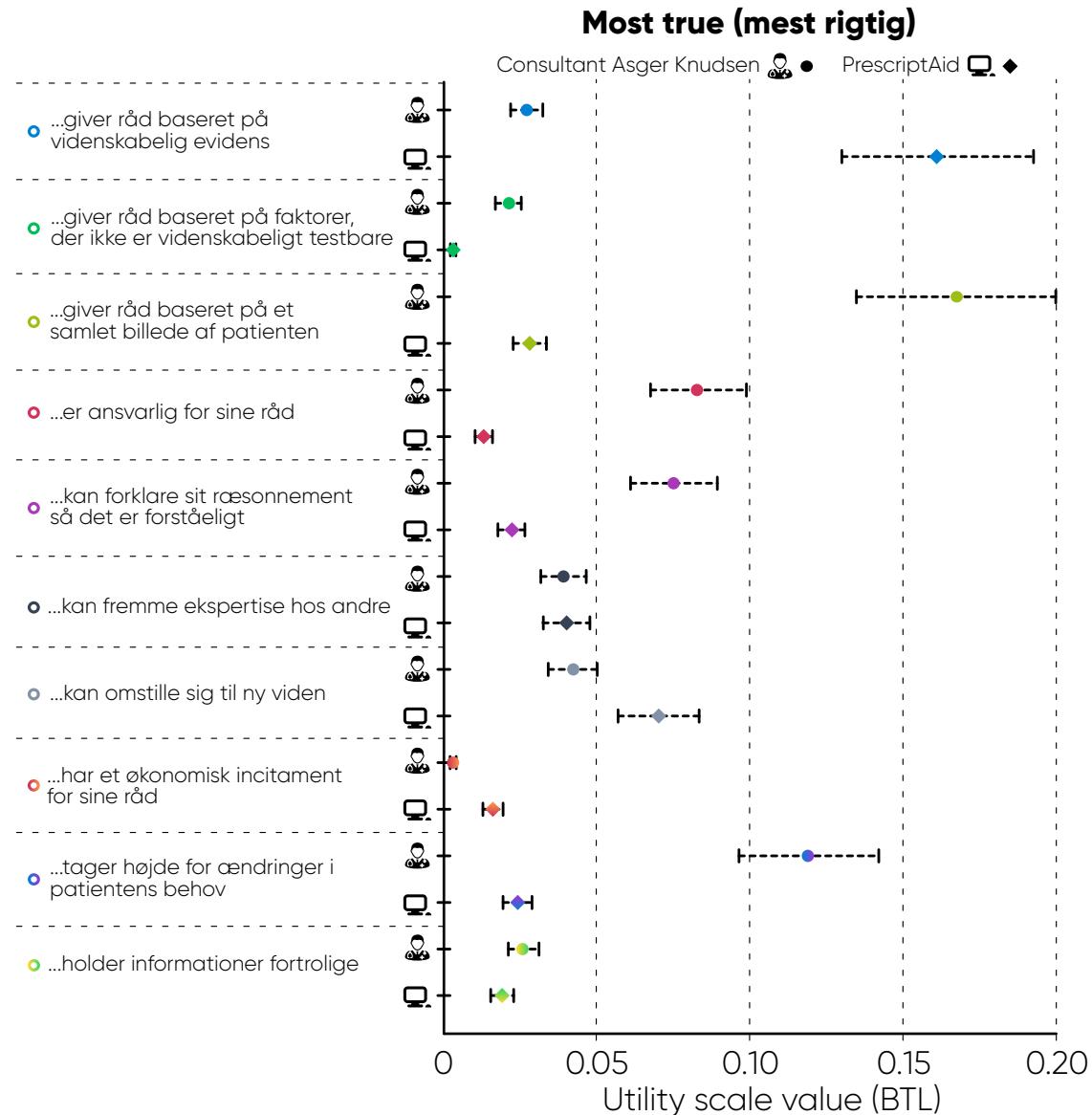


Figure 5.5: Shows normalised ratio scale values from the BTL model as well as corresponding confidence intervals. PrescriptAid is represented by diamonds and Consultant Asger Knudsen as dots. The statements are represented by the colours described above.

Table 5.4: Shows the normalised BTL model scale values (U scale), their corresponding 95% confidence intervals, and a rank ordering of the stimuli with the statement perceived as most true is 1 while the statement perceived as the least true is 20.

Label	Char.	Stimuli	U scale(95% CI)	Rank
A		giver råd baseret på videnskabelig evidens.	0.0271(0.0219,0.0324)	11
B		giver råd baseret på videnskabelig evidens.	0.1611(0.1298,0.1924)	2
C		giver råd baseret på faktorer der ikke er videnskabeligt testbare.	0.0214(0.0171,0.0256)	15
D		giver råd baseret på faktorer der ikke er videnskabeligt testbare.	0.0029(0.0020,0.0038)	20
E		giver råd baseret på et samlet billede af patienten.	0.1673(0.1348,0.1998)	1
F		giver råd baseret på et samlet billede af patienten.	0.0281(0.0227,0.0335)	10
G		er ansvarlig for sine råd.	0.0826(0.0670,0.0983)	4
H		er ansvarlig for sine råd.	0.0135(0.0106,0.0163)	18
I		kan forklare sit ræsonnement så det er forståeligt.	0.0750(0.0609,0.0892)	5
J		kan forklare sit ræsonnement så det er forståeligt.	0.0223(0.0179,0.0267)	14
K		kan fremme ekspertise hos andre.	0.0391(0.0317,0.0465)	9
L		kan fremme ekspertise hos andre.	0.0402(0.0326,0.0477)	8
M		kan omstille sig til ny viden.	0.0423(0.0343,0.0503)	7
N		kan omstille sig til ny viden.	0.0702(0.0570,0.0834)	6
O		har et økonomisk incitament for sine råd.	0.0034(0.0024,0.0044)	19
P		har et økonomisk incitament for sine råd.	0.0160(0.0128,0.0193)	17
Q		tager højde for ændringer i patientens behov.	0.1188(0.0960,0.1416)	3
R		tager højde for ændringer i patientens behov.	0.0242(0.0195,0.0289)	13
S		holder informationer fortrolige.	0.0257(0.0207,0.0307)	12
T		holder informationer fortrolige.	0.0186(0.0149,0.0224)	16

As can be seen on [Figure 5.5](#) and [Table 5.4](#), the confidence intervals are of varying sizes, allowing for some analysis into differences between the statements. Since the results are ratio scale values, it is possible to interpret how much more true the statements are perceived relatively to each other. By comparing the same statements for both the consultant and the decision aid, it is possible to analyse the differences between the statements for the two advisers.

It is seen that the same statement is perceived as more true for the decision aid than for the consultant in three instances. The first statement (...giver råd baseret på videnskabelig evidens) is perceived to be 4 (4.006) to 9 (8.785) times more true for the decision aid than for the consultant. The second statement (...kan omstille sig til ny viden) is perceived to be 1.1 (1.133) to 2.5 (2.431) times more true for the decision aid than for the consultant. The third statement (...har et økonomisk incitament for sine råd) is perceived to be 3 (2.909) to 8 (8.042) times more true for the decision aid than for the consultant.

It is seen that the same statement is perceived as more true for the consultant than for the decision aid in five instances. The first statement (...giver råd baseret på faktorer, der ikke er videnskabeligt testbare) is perceived to be 4.5 to 12.8 times more true for the consultant than for the decision aid. The second statement (...giver råd baseret på et samlet billede af patienten) is perceived to be 4 (4.024) to 9 (8.802) times more true for the consultant than for the decision aid. The third statement (...er ansvarlig for sine råd) is perceived to be 4 (4.110) to 9 (9.274) times more true for the consultant than for the decision aid. The fourth statement (...kan forklare sit ræsonnement, så det er forståeligt) is perceived to be 2 (2.281) to 5 (4.983) times more true for the consultant than for the decision aid. The fifth statement (...tager højde for ændringer i patientens behov) is perceived to be 3 (3.322) to 7 (7.262) times more true for the consultant than for the decision aid.

The remaining two statements (...kan fremme ekspertise hos andre and ...holder informationer fortrolige) have overlapping confidence intervals, and are therefore seen to not likely be differing.

Furthermore, to better compare the statements a general rank ordering of the statements can be seen in [Table 5.4](#), this contains groupings of the statements with overlapping confidence intervals, which can also be seen on [Figure 5.5](#). The combinations are visualised on [Figure 5.6](#).



Figure 5.6: Shows groups of statements with overlapping confidence intervals for all statements, decision aid only and consultant only. The further to the top the more the statement is perceived to be true.

[Figure 5.6](#) does not show how true a statement is perceived compared to another, it does however give a brief overview of the rank ordering of the statements as based on the BTL model as well as which statements might be seen as the most true. By comparing the rank ordering from [Figure 5.6](#) with [Figure 5.5](#) and the values from [Table 5.4](#) it is possible to analyse relevant differences in the data.

For the decision aid only, it is seen that the statement B (giver råd baseret på videnskabelig evidens) is perceived as more true than the rest of the statements. Statement B is perceived as

1.5 (1.556) to 3.4 (3.375) times more true than statement N (kan omstille sig til ny viden), which is perceived as more true than the remaining statements for the decision aid. Statement N is perceived as 1.2 (1.195) to 2.6 (2.558) times more true than statement L (kan fremme ekspertise hos andre). The statement perceived as being the least true is D (giver råd baseret på faktorer der ikke er videnskabeligt testbare), which is perceived to be 2.8 (2.789) to 8.15 times less true than statement H (er ansvarlig for sine råd) and 34 (34.026) to 962 times less true than statement B which was perceived as most true.

For the consultant the statement perceived to be most true is E (giver rår baseret på et samlet billede af patienten), which has overlapping confidence intervals with the next statement Q (tager højde for ændringer i patientens behov), it is therefore not possible to determine whether these statements actually differ from each other. E is perceived to be 1.4 (1.371) to 3 (2.982) times more true than statement G (er ansvarlig for sine råd). An interesting divide is between statement I (kan forklare sit ræsonnement så det er forståeligt) and M (kan omstille sig til ny viden), since M does not have overlapping confidence intervals with the four statements perceived as more true. I is perceived as 1.2 (1.196) to 2.6 (2.601) times more true than M and E is perceived to be 2.7 (2.680) to 5.8 (5.825) times more true than M. An interesting find is that A (giver råd baseret på videnskabelig evidens) does not overlap with M and is perceived as 1.06 (1.059) to 2.3 (2.297) times less true. The statement perceived as least true is O (har et økonomisk incitament for sine råd), which is perceived as 3.886 to 10,667 times less true than C (giver råd baseret på faktorer der ikke er videnskabeligt testbare) and 35,5 (35,474) to 832,5 times less true than E which was perceived as most true. Another interesting find is that statements A and C does have overlapping confidence intervals.

Another way to visualise the data is by checking for additional common attributes between the stimuli. This is done by fitting an EBA model, which can take the shape of a preference tree.

5.4.3 EBA model (preference tree)

Even though the BTL model indicates that the underlying perceived correctness of the statements is uni-dimensional, it is checked whether a preference tree (EBA) model might have a better fit to the statistical model. This is used to study whether an underlying hierarchically structure of additional common attributes provides a better fit. Different preference tree structures including a structure based on "Human vs Computer" ($\chi^2(169) = 225.2, p = 0.00251$) and a structure based on "Most picked" ($\chi^2(169) = 593.8, p = < 2 * 10^{-16}$) were tested, but neither provided a better fit. However, when testing a structure based off of which statements describes similar aspects the model provided a fit as good as the one described by the BTL model. However, the best fit ($\chi^2(169) = 142.2, p = 0.934$) was found as seen in [Figure 5.7](#) by using the following combination: (A = [1], [2], [3,22], [4], c(5,21), [6,21], [7], [8], [9,22], [10], [11], [12], [13], [14], [15,21], [16,21], [17], [18], [19], [20]). As the BTL model can be regarded as a preference tree without branches (additional common attributes), it is possible to compare the two models statistically by a likelihood ratio test. This shows that the preference tree model creates a significantly better fit than the BTL model ($\chi^2(2) = 8.6477, p = 0.01325$). The labels used in [Figure 5.7](#) can be seen in [Table 5.5](#) along with scale values of the preference tree. The "U" scale shows the full length of the different stimuli. The full length of the stimuli that is branched out into either "21" or "22", is found by combining the attribute (A) scale value of "21" or "22" with the attribute scale of the specific stimuli.

All the attribute scales seen in [Table 5.5](#) can be seen in [Figure 5.8](#). Here a visual representation of the attribute scales are plotted with their corresponding standard error to the mean.

This preference tree model was obtained by assuming that only two branches were present, with one separating stimuli "I", "J", "K" and "L" and one separating stimuli "E" and "Q" from the rest. The stimuli in the first branch include both computer and human stimuli present, while the second only consisted of human stimuli. The first branch describes is found to describe communication skills, both in terms of enhancing expertise in others, but also being able to explain their reasoning. The second branch describes being able to understand the patient in question, both in terms of understanding all factors that affect the patient and when a change in the needs of a patient occurs.

Table 5.5: Shows the preference tree labels, U scale value for each stimuli, all attribute scale values and standard error values.

Label	Char.	Stimuli	U scale	Attr. nr.	A Scale	SE
A		giver råd baseret på videnskabelig evidens.	0.0328	1	0.0328	0.0034
B		giver råd baseret på videnskabelig evidens.	0.1935	11	0.1935	0.0195
C		giver råd baseret på faktorer der ikke er videnskabeligt testbare.	0.0258	2	0.0258	0.0027
D		giver råd baseret på faktorer der ikke er videnskabeligt testbare.	0.0035	12	0.0035	0.0006
E		giver råd baseret på et samlet billede af patienten.	0.1848	3	0.0425	0.0389
F		giver råd baseret på et samlet billede af patienten.	0.0339	13	0.0339	0.0035
G		er ansvarlig for sine råd.	0.0995	4	0.0995	0.0100
H		er ansvarlig for sine råd.	0.0163	14	0.0163	0.0018
I		kan forklare sit ræsonnement så det er forståeligt.	0.0858	5	0.0706	0.0114
J		kan forklare sit ræsonnement så det er forståeligt.	0.0285	15	0.0133	0.0053
K		kan fremme ekspertise hos andre.	0.0473	6	0.0320	0.0074
L		kan fremme ekspertise hos andre.	0.0479	16	0.0327	0.0077
M		kan omstille sig til ny viden.	0.0510	7	0.0510	0.0052
N		kan omstille sig til ny viden.	0.0845	17	0.0845	0.0085
O		har et økonomisk incitament for sine råd.	0.0042	8	0.0042	0.0006
P		har et økonomisk incitament for sine råd.	0.0194	18	0.0194	0.0021
Q		tager højde for ændringer i patientens behov.	0.1540	9	0.0116	0.0127
R		tager højde for ændringer i patientens behov.	0.0292	19	0.0292	0.0031
S		holder informationer fortrolige.	0.0311	10	0.0311	0.0032
T		holder informationer fortrolige.	0.0225	20	0.0225	0.0024
				21	0.0152	0.0071
				22	0.1424	0.0299

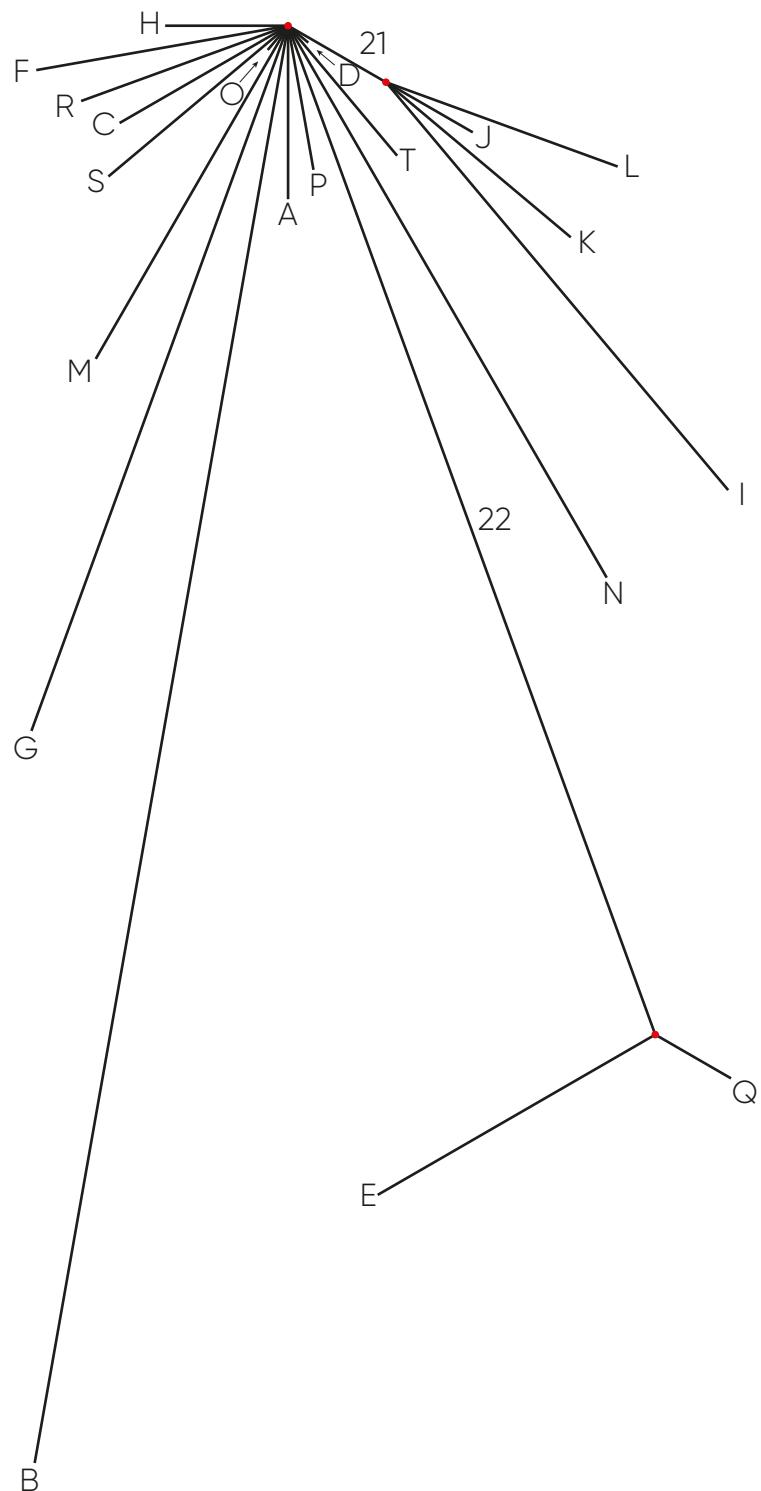


Figure 5.7: Shows the preference tree with stimuli A-T and attributes 1-22. For specific scale values see [Table 5.5](#)

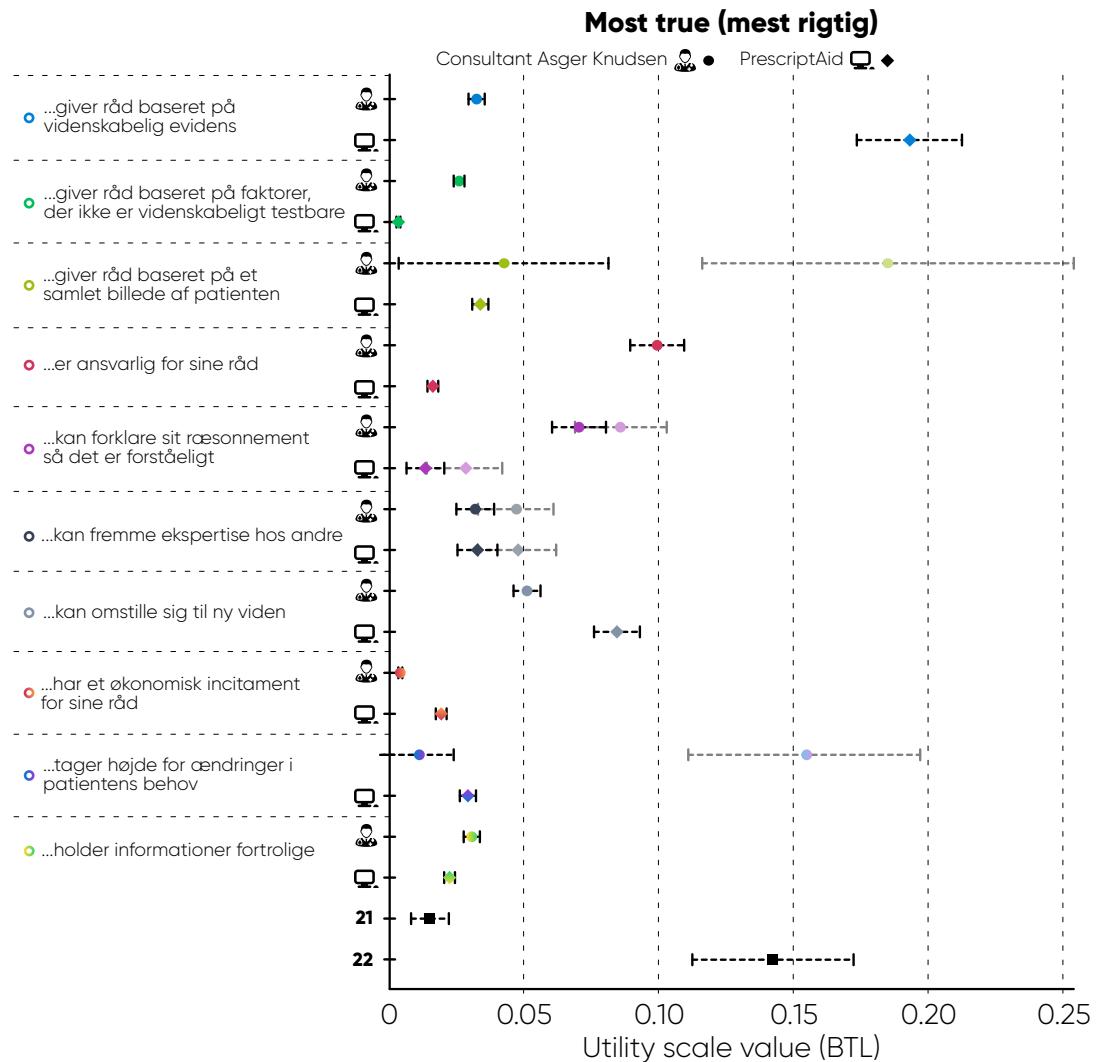


Figure 5.8: Shows A scale values from the EBA (preference tree) model as well as corresponding standard errors. PrescriptAid is represented by diamonds and Consultant Asger Knudsen as dots. The additional A scale values are represented by squares. The statements are represented by the colours described above. The U scale values that differ from the A scale values are represented at half opacity.

This chapter has described the experimental study as conducted in this thesis as well as the results gathered from the participants. Furthermore, the data analysis has found a BTL model with a good fit and a EBA (preference tree) model with a better fit. How the results can be used to answer the research questions is described in the following chapter. The focus is on how healthcare professionals perceive other healthcare professionals and decision aids, as well as how the probabilistic choice models are used to create decision aids that are trusted by healthcare professionals.

6

General discussion

This chapter contains a general discussion of the results as presented in section 4.4 and section 5.4. The starting point of the discussion is the three research questions as presented in chapter 3. Furthermore, the methods used during the thesis are discussed with the purpose of shedding light on advantages and disadvantages as well as the sources of error and how these might affect the results and what might be interpreted from these. Lastly the discussion touches upon future work relevant to pursue after the conclusion of the study.

6.1 Healthcare professionals and trust in decision aids

The first research question seeks to explore whether healthcare professionals trust or distrust decision aids and algorithmic advice, to what extent they trust or distrust decision aids and why. The research question is as follows:

Do healthcare professionals distrust decision aids and algorithmic advice, and why?

As previously described, there are studies into people's trust in decision aids (algorithms and algorithmic advice) (Dietvorst et al., 2015, 2016; Logg et al., 2019; Önkal et al., 2009; Prahla & Van Swol, 2017; Promberger & Baron, 2006), few have studied the effects of expert professionals and their trust in decision aids (Logg et al., 2019). Furthermore, these expert professionals have not been healthcare professionals. The newest research challenges the assertion that people are averse to decision aids and algorithmic advice and shows that people have a general trust towards decision aids and algorithmic advice (Dietvorst et al., 2015, 2016; Logg et al., 2019; Prahla & Van Swol, 2017). Expert professionals on the other hand are seen to have a distrust to any advice, decision aid or other people. This research is based on experimental *Weight of Advice* studies in which participants are tasked to make a guess on a given subject, e.g. the weight of a person or how well a song might perform on the hit-lists. These situations are far from the real life situations in which decision aids are involved, which gives a low ecological validity of the results. The research does however have a strong internal validity due to the controlled environment of the studies as well as direct measurements of how much weight participants put on advice from different advisers.

Our research explored healthcare professionals and their trust in decision aids through semi-structured and contextual interviews. The exploratory study generally did not find a distrust towards decision aids in the Danish healthcare sector, it did however find that the Regions of Denmark and the Danish government are investing in new decision aids for the healthcare sector. This however mainly shows a trust of decision aids in the administration of the healthcare sectors and hospitals as well as the IT departments. The main advantages as seen from their point of view are the efficiency improvements that can be made to provide cheaper healthcare as well as improvements to patient security to provide a safer and more trustworthy healthcare sector. The administration and IT departments does however only represent two (administration and IT) out of four pillars in the Danish hospitals. Therefore they do not represent the nurses and the clinicians of the Danish hospitals. Neither do they represent general practitioners or private hospitals in Denmark. During the exploratory study it was found that healthcare professionals did not have the same amount of optimism towards decision aids being implemented in healthcare, a general trust towards decision aids and their advantages was however found.

As part of exploring healthcare professionals' trust towards decision aids a list of 23 statements describing attributes relevant for healthcare professionals to trust decision aids was found. The statements generally put a large emphasis on being able to understand context, being able to give advice based intuition and on parameters not scientifically measurable, being able to adapt to new information, being grounded in science and scientifically approved, show interest in patients, keep information confidential, know critical from non-critical situations, being able to explain reasoning, origin of information, and being able to explain the processes

underwent for the given advice as well as being responsible for the advice and being economically independent. The statements were grouped based on relevance to each other and how similar they were, this resulted in the 10 statements used in the following experimental study.

The experimental study was performed in order to measure how healthcare professionals perceive the 10 statements to be true for a decision aid, in turn creating an overview of healthcare professionals trust in a decision aid as measured on the basis of these statements. The experimental study is able to show that healthcare professionals and medicine students perceive the decision aid to be able to give advice based on scientific evidence and that the decision aid is able to adapt to new knowledge to a much larger extent than they perceive the remaining statements to be true. Furthermore, the decision aid is perceived as not being able to give advice based on parameters that are not scientifically testable. Interestingly the decision aid is perceived as being able to enhance the expertise of others to a larger extend than being able to explain its reasoning in an understandable way. The decision aid is not perceived to be able to keep information confidential as compared to the four most true statements, although to a larger extend than giving advice that is not scientifically testable. Furthermore, the decision aid is not perceived as being responsible for the advice given and is seen to not have an economic incentive for its advice, although still to a larger extent than giving advice that is not scientifically testable. Since the ratio scale values for the statements are relative to each other it is relevant to compare the results for the decision aid to the results of the consultant.

6.2 Trust in advice from other healthcare professionals

The second research question seeks to explore whether healthcare professionals trust or distrust other healthcare professionals and their advice, to what extent they trust or distrust other healthcare professionals and why. The research question is as follows:

Do healthcare professionals trust advice from other healthcare professionals, and why?

As previously described Logg et al. (2019), suggest that expert professionals not only have a distrust in algorithmic advice from decision aids, but also a general distrust to advice from all sources. This is as the previously mentioned studies based on an experimental weight of advice study, in which the experts make a guess themselves, get a suggestion from an algorithm (or other people) and has the possibility to alter their initial guess based on this. The expert professionals tend to alter their initial advice less than laypeople. Another way to interpret this effect could be that expert professionals tend to weigh their own advice above algorithms' (and others') in their field of expertise, which does not necessarily entail a distrust in others or algorithms.

That healthcare professionals trust each other, was also found during the exploratory study conducted. During interviews it was found that healthcare professionals did not just trust other healthcare professionals with their own background, but that they were quite willing to take advice from each other despite their different backgrounds and work roles. It was found that the Danish healthcare sector is perceived as flat hierarchically with easy exchanges of information across entire departments. Some notes were however made during exploratory interviews and exit interviews about learning to recognise bright colleagues that are more likely to correct them when they make mistakes and give good advice.

The experimental study created an overview of how healthcare professionals perceive other healthcare professionals, more specifically how they perceive a 64 year old consultant (ovrlæge) as described in Appendix G. The consultant is perceived as being able to give advice based on a complete image of the patient and being able to take changes in the needs of patients into account. Furthermore, the consultant is perceived as being responsible for their advice and being able to explain their reasoning compared to the remaining statements. The consultant is not perceived to have an economical incentive compared to the remaining statement. Interestingly giving advice based on scientific evidence and giving advice based on factors not scientifically testable did have overlapping confidence intervals, meaning it is likely not possible to find a significant difference in how true these statements are perceived to be compared to each other. The ratio scale values for the statements explaining how true they are perceived by the participants, can be compared for the consultant and the decision aid in order to gain a deeper insight into how these might be perceived differently by other healthcare professionals.

6.3 Perception of advice from a consultant and a decision aid

Studies into people's trust in decision aids and expert professionals (Önkal et al., 2009; Promberger & Baron, 2006), shows that people will rather trust human expert professionals than they will decision aids and forecasting algorithms. The study by Önkäl et al. (2009) is carried out as a weight of advice experiment and has the advantages and disadvantages as previously described studies. Promberger and Baron (2006) shows that participants would rather follow recommendations made by a physician than recommendations made by a computer. Participants were instructed to choose whether they would like to follow the instructions provided by the adviser or not. As compared to the weight of advice studies the results might seem more ecologically valid, although the experiment was not performed on actual patients in hospitals, but by people assigned to an online survey. Furthermore, the purpose of the study may be obvious to the participants which might affect the results. Common for both studies is that they involve laypeople and their trust in human experts versus decision aids. As previously mentioned Logg et al. (2019) found that expert professionals were not likely to trust advice from any source, human or decision aid. This although did not look at expert professionals trust in other expert professionals, although a likely effect was suggested.

Interestingly, the statements found during the exploratory study are largely similar to the anecdotal explanations of the causes of people not trusting advice from algorithms. Namely, the ability for algorithms to learn as suggested by Dawes (1979) and to improve through experience as suggested by Highhouse (2008) and the found statement about being able to adapt to new knowledge. Furthermore, the notion that algorithms cannot incorporate qualitative data (Grove & Meehl, 1996) and are dehumanising (Dawes, 1979; Grove & Meehl, 1996) are described with similar meanings as is the basis for the statements about being able to take into account the needs of the patient and provide advice based on a complete picture of the patient. The suggestion of an effect based on the need for algorithms to be perfect (Dawes, 1979; Einhorn, 1986; Highhouse, 2008) are similar to the statement of the algorithm to give advice based on scientific evidence. Lastly, the suggestion that not wanting to rely on algorithms for important decisions might affect trust in them (Dawes, 1979; Shariff et al., 2017) is similar to the statement about having responsibility for the given advice.

With the statements being of similar character, it could be possible to rethink the explanations for algorithm aversion. The statement that the algorithms need to be perfect corresponds well with the statement that algorithms are able to give advice based on scientific evidence. This statement was perceived as being the most true for the decision aid and also found to be more true than the same statement for the consultant, which might suggest that this effect is not as important as suggested by Dawes (1979); Einhorn (1986); Highhouse (2008) for healthcare professionals. The remaining anecdotal explanations seem to correspond with the results found in the experimental study.

The experimental study shows that the results from the individual participants are consistent. Furthermore, it is found that the results show few stochastic transitivity violations and that it is possible to fit both BTL and EBA (preference tree) models with good fits ($p>0.1$). This suggests that hypothesis III is rejected, since subjects are able to consistently choose between the statements.

The experimental study of this thesis shows a range of differences between the participants' perception of a consultant and a decision aid. As mentioned previously three statements for the decision aid is perceived to be more true than the same statement for the consultant. It is shown that there is a general trust for the decision aid to provide advice based on scientific evidence compared to the consultant with a factor of about four to nine. The decision aid is found to be perceived as being able to adapt to new knowledge compared to the consultant, this is found to be 1.1 to 2.5 times more true. It is also found to be more true that the decision aid has an economical incentive when it provides advice. This confirms hypothesis II for these statements while rejecting hypothesis I. Therefore, it is found that healthcare professionals and medicine students trust the ability for the decision aid to provide advice based on scientific evidence to a larger extent than the consultant. The decision aid is trusted to be able to adapt to new knowledge to a larger extent than the consultant. Furthermore, it is found that the decision aid to a larger extent is trusted to have an economical incentive for its advice. It is however unclear whether this is due to the decision aid being programmed with larger economical scopes in mind by the government and regions or it is seen as having the possibility to be used by the medical industry to recommend specific medicine and enhance sales. Further study into the nuances of this statement is needed in order to make conclusions.

A total of 5 statements are found to be more true for the consultant than the same statements are for the decision aid. This confirms hypothesis I for these statements while it rejects hypothesis II. These include giving advice based on factors that are not scientifically testable,

that he gives advice based on a complete picture of his patient, is responsible for his advice, is able to explain his reasoning in a way that is understandable, and takes changes in the needs of the patient into account. Therefore, it is found that healthcare professionals and medicine students trust the consultant to provide advice that are based on a complete picture of their patient, that takes into account the needs of the patient, and is based on factors not necessarily scientifically testable to a larger extent than the decision aid. The consultant is also trusted to be able to explain their reasoning in an understandable way and is responsible for his advice to a larger extent than the decision aid.

The largest differences between the statements are that the decision aid is perceived to provide advice based on scientific evidence compared to the consultant (4 to 8.8 times), that the decision aid has an economical incentive (2.9 to 8 times), that the consultant is perceived to provide advice based on complete picture of his patient (4.5 to 12.8 times), that the consultant is responsible for his advice (4.1 to 9.3 times), that the consultant is able to explain his reasoning in an understandable way (2.3 to 5 times), and that he takes changes to the patients' needs into account (3.3 to 7.3 times).

The remaining two statements about being able to enhance the expertise of others and being able to keep information confidential has overlapping confidence intervals and it is not possible to determine whether there is a difference between these for the two advisers. This rejects both hypothesis I and II as a difference between the two advisers is not likely to be found. Therefore, it is found that the healthcare professionals and medicine student do not trust decision aids ability to enhance their expertise and keep information confidential any less than they do a medical consultant.

As described, it is not possible to confirm or reject either hypothesis I or II for all statements. Hence, it is found that trusting one adviser more than the other, is dependent on multiple factors as described by the statements.

6.4 Improving decision aids based on probabilistic choice models

The third research question seeks to explore whether the results gather throughout the exploratory study and experimental study in the form of probabilistic choice models can be used to optimise healthcare professionals' trust in decision aids. The research question is as follows:

How are decision aids designed to optimise trust in the given advice?

According to Dietvorst et al. (2015), seeing algorithms err decreases people's trust in advice from algorithms (decision aids). Furthermore, it is shown that increasing the user's (or advice receiver's) possibilities of modifying the advice increases the trust in the advice given by algorithms (decision aids). Furthermore, Logg et al. (2019) found that the trust in algorithmic advice increased with the numeracy of the participants. A possible suggestion to the cause is the lack of understanding of what goes on 'inside' of the algorithm. If this is the case, it is relevant to include the work of Gigerenzer and Edwards (2003) in order to best optimise advice to the societal statistical illiteracy described by him.

To optimise the trust of decision aids, it is relevant to enhance the statements perceived to be less true for the decision aid compared to the consultant. Hence, it is useful to improve healthcare professionals' perception of decision aids on these factors.

6.4.1 Giving the complete picture of the patient

To optimise trust in the decision aids ability to give advice based on a complete picture of the patient, it is possible to put larger emphasis on, which data the decision aid is able to 'see'. This could be done by visualising the input data of the decision aid. This could include key measurements about the condition of the patient, e.g. updated blood tests, weight, and other relevant data. It is found to be important that these data are visualised in a way, that the healthcare professionals using the system, are able to view them when necessary, possibly providing the most relevant data at all times. This functionality was not included in the decision aid description for the experimental study. Furthermore, it could be useful to visualise the possibility of values commonly believed to be immeasurable or values measured but commonly not thought of as such, such as galvanic skin response, emotion recognition using computer vision, etc. This could have the effect of enhancing the healthcare professionals knowledge about what is in fact measurable by modern sensor technologies.

Another way to emphasise that the decision aid is giving advice based on a complete picture of the patients could be to put the patient into the centre of the systems interface. By showing a

human patient, with the most important values highlighted could provide a trust in the decision aids ability to take into account the whole patient.

6.4.2 Take changes to the needs of the patient into account

To enhance healthcare professional's trust in the decision aid's ability to give advice that take the needs of the patient into account, it is possible to build upon the changes suggested to enhance trust in advice from a complete picture of the patient. Putting emphasis on the availability and updateability of the data used by the system in order to provide advice. This could include emphasis on when data were last updated, and emphasis on technologies that can be used/developed in order to provide an even easier transfer between healthcare devices and the decision aid system. During the contextual interview, it was found that the current devices used for blood pressure, weight, height, and width measurements did not transfer data to their systems automatically and it is necessary to manual input the data. Therefore, highlighting possible solutions for the automatic transfer of data between devices could be a way for the decision aid to better take into account changes in the needs of the patients. Furthermore, the decision aid could suggest when new values might be necessary, for instance outdated blood pressure measurements, blood test, and blood sugar levels, to provide a trust in the decision aids ability to be context aware and 'intelligent'. A way for the decision aid to suggest when values are critical or when values may be outliers and a new test might be useful could also provide the healthcare professional with information that support the patient and takes changes to the needs of the patient into account.

Since the decision aid is found to be perceived at being able to adapt to new knowledge compared to the consultant, it is suggested that the healthcare professionals are in fact able to understand the origin of the data used for giving advice that are not dependent on the patient and the internal devices and policies of the hospitals.

6.4.3 Being able to visualise the reasoning in an understandable way

To make the reasoning that are the basis of advice given by the decision aid, further information could be provided to support the explanations to a larger extent. The reasoning behind the advice of the decision aid were already present in the current description of the decision aid. Hence it is relevant to study how these can be altered in order to further emphasise the origin and explanations of advice as well as how these can be presented in a way such that they are more understandable. Furthermore, better data visualisation using spider plots, graphs, and simulations of effects of different scenarios and choices that can be made, can be implemented in order to support the advice given by the decision aid.

6.4.4 Scientific versus non-scientific

It is seen that the decision aid is perceived to be able to provide advice based on scientific evidence to a larger extent than the consultant. Furthermore, it is seen that the consultant is perceived as being able to provide advice based on factors that are not scientifically provable. Factors not scientifically testable is understood as factors that are difficult or not able to be tested based on the quite narrow scientific understanding of scientific evidence found among clinicians, where randomised control studies on large subject groups and meta analysis of these with a large focus on significance testing is the focus. It is arguable whether the ability to take into account information that might be perceived as not scientifically provable is an advantage or a disadvantage. According to the contextual interview the interviewed consultant put much emphasis on the fact that he is able to trust his gut feeling and that this is not achievable by any decision aid. Algorithm and decision aid researchers on the other hand argues that algorithmic advice from decision aids are better even when considering human gut feeling (Harell, 2016; Logg, Minson, & Moore, 2018).

To improve trust in the decision aid according to the perception of being able to give advice based on factors that are traditionally not seen as scientifically testable, it is important to highlight the basis of the data used for the piece of advice given. This could be done as previously mentioned by better visualisations and more transparency in data origin. Furthermore, highlighting measurements of the patient that could possibly support better advice could be implemented such as suggestions to use computer vision, skin responses and possibly standardised questionnaires about the patients' experiences and symptoms.

6.4.5 The role of economical incentive in healthcare

Interestingly, even though the consultant is perceived as giving advice based on factors not scientifically testable, he is perceived as having less economical incentive compared to the decision aid. This could support the suggestion that economical incentive is understood as societal economic incentive that seeks to give the most economically responsible advice as seen from a societal point of view. In order to highlight whether the decision aid is perceived as having an economical incentive, data origin could be highlighted when advice are given. Furthermore, it is possible to clearly state, if advice are considered an economical advantage for the society as a whole. To support that the decision aid does not have an economical incentive towards the medicine industry, it is important be transparent in the data used for the recommendations made by the decision aid. It is furthermore, important to make sure data used are as objective as possible and does not include ties to the medical industry.

The contextual interview included an example of economic incentive about what time of day patients should be released from the hospital. This could include being aware that some patients need to travel with a ferry to get home, and hence will need to be released in good time in order to not be forced into an extra night at the hospital with economical consequences for the department. Hence, it is important to include data that are not often considered scientific data about the health of the patients. Including relevant data such as these might increase the perception that the decision aid has economical incentive as well as is able to give advice based on factors that are not scientifically testable, a complete picture of the patient, and take into account changes in the patient's needs. Even though the consultant might be able to take into account whether their patients need to be released in order to go with a ferry, they are often not aware of these factors. Hence, this could be the explanation for why the consultants are not perceived as having an economical incentive.

Another explanation could be that consultant are generally perceived with the main purpose of taking care of the health of the patient, while the decision aid may be perceived as an economical investment from the Danish Regions with the purpose of streamlining the Danish healthcare by making it more efficient. This explanation is backed by the contextual interview in which it was found that implemented healthcare systems were seen as controlling and that they were removing the competencies of the healthcare professionals with the purpose of making healthcare more efficient and less costly. Hence, a common understanding among healthcare professionals that the main purpose of healthcare systems have been to improve efficiency rather than for the improve the conditions of the patients.

6.4.6 Being responsible for the given advice

The last statement perceived as being more true for the consultant than the decision aid is being responsible for the given advice. This may be the most difficult subject to design for, since this is could be considered an ongoing ethical discussion of who are responsible for advice given by decision aids and choices made by algorithms such as self-driving cars. Decisions and advice that may cause damage to property or even risk harming or threatening human lives. During interview II it was found that current machine learning based algorithms and decision aids for the Danish healthcare is designed in a way that decisions are always made by clinicians. This works inside of the current societal norms, in which people are responsible for their actions even though they might have been influenced along their decision making processes. According to Interview I, it was found that clinicians and other healthcare professionals rarely get charged with being responsible for unintended events happening to patients. According to interview I it can, however have consequences for the involved healthcare professionals. In situations where the healthcare professionals intend to cause harm, punishments are much more likely to occur.

A way to ensure proper responsibility for the advice given, may be to clearly represent where data originate from and log who and when changes to these data are made. This could be the case for data received from external sources such as *pro.medicin*, *interaktionsdatabasen*, *maxdosis*, or *Cave* as was used in the decision aid used here and planned for implementation across Denmark. It is also relevant to log and visualise from where data collected by the system originates in order to be able to spot unintended events early on or to be able to change collection procedures and avoid future unintended events. To follow the current procedures and culture in Danish healthcare, it is important to avoid the ability to punish individuals for unintended events, unless they are done purposefully and thus are not unintended.

Another aspect of responsibility lies in keeping information confidential. The consultant and decision aid are both trusted to keep information confidential less than most factors of the experimental study. Hence, is is relevant to ensure trust in the data being confidential. During the contextual interview, it was found that informed consent is a big part of patient security

of their information and data. It was however also found that this is difficult to ensure with IT systems since they are not always able to ask for informed consent before giving advice or gathering information. Furthermore, hacking of healthcare systems was mentioned briefly as a concern. Hence, it is important to ensure security of the data kept and processed by the decision aid. Visualising this to the healthcare professionals could support their perception of systems being able to keep information confidential.

6.4.7 Enhancing the expertise of others

It is seen that healthcare professionals seem to perceive the consultant and decision aid to both be able to enhance the expertise of others to the same extent. Making changes to the decision aid as described in the previous subsections may cause the decision aid to be perceived as being able to enhance expertise of others to a larger extent than the consultant.

6.4.8 Additional attributes from the preference tree model

As reported in subsection 5.4.3 a better fit was found for a preference tree model. Overall the preference tree model was not found to differ much from the BTL model, but two branches indicate that some stimuli may have additional common attributes. The first of the two, consisted of statements related to communication skills. This could possibly show that communication skills could be an additional attribute, and using this as an overall factor when designing decision aids would benefit the underlying attributes. The second branch could describe an ability to understand the patient and their needs, here only the human consultant versions of the statements were grouped which could possibly show an attribute related to human understanding or compassion. Even though the preference tree model had a better goodness of fit than the BTL, not many additional conclusions can be made from the model.

6.5 Evaluation of chosen methods

This section contains a discussion of the chosen methods, how they could be improved upon, and the advantages and disadvantages of this.

6.5.1 Qualitative interview studies

The exploratory study with the purpose of gathering information about the Danish healthcare sector and healthcare professionals' perceptions of decision aids and receiving advice. It is found that experts provided good information to be the basis of the experimental study. Furthermore, it would have been beneficial to carry out additional contextual interviews with clinicians to get a broader and more filling description of the context in which they use decision aids. This could possibly have been done with a broader range of clinicians conducting activities and tasks with the help from a decision aid. This could have provided a broader foundation for the character descriptions, the scenario, and possibly provided additional statements for the experimental study. It was however found that the interview study, including semi-structured and contextual interview, reached a point of saturation in which not much new information was gathered about the perceptions of decision aids with new interviews.

6.5.2 Sorting and reduction of statements

A disadvantage of this study is that the delimitation of statements describing healthcare professionals was made by the authors. An advantageous approach could have been to have healthcare professionals sort the statements into groups of similar statements. This could have provided the basis of the delimitation. Another way of reducing the amount of statements could have been to perform multi dimensional scaling or factor analysis on the basis of a rating or similarity experiment. Hence, a simple pairwise comparison experiment in which healthcare professional can rate statements according to how similar they are. This experiment would be performed without any adviser characters in mind and with the purpose of mapping the similarity of the statements in order to reduce the amount of statements needed in the pairwise comparison study including the two advisers. This could have provided a better reduction and delimitation of the amount of statements grounded in the perceptions of healthcare professionals.

6.5.3 Evaluation of the pairwise comparison study

To evaluate the pairwise comparison study conducted during this thesis, the sources of error are discussed and compared to the advantages of the method. Furthermore, the validity of the results and the cautiousness that should be had, when making conclusions and assumptions based on the results, due to the sources of error and the disadvantages of the pairwise comparison method.

Sources of error

As seen in the exit interview in [section L.3](#), six of the participants expressed the fact that they felt that they changed attitude towards some of the statements presented for them throughout the experimental study. The participants reported that halfway through they either changed their view on whether a statement was true or not, or how they understood said statement. This could likely lead to inconsistency in the data collected, but as explained in [subsection 5.4.1](#) this was not the case. Based on the experimental setup it could be that the participants feel that they are inconsistent because they received many of the statements multiple times and are not totally sure on what answer they gave the last time. As described the experimental setup was done so the participant would not receive the same two statements again, but the large amount of total comparisons could lead to them feeling that they did, and that they picked a different answer.

There could however still be a potential effect of the participant changing how they understood the statements, and this could possibly have been avoided with either describing the statements in more detail, changing the statements to create an easier understanding or starting the experiment with a familiarisation round. The familiarisation could consist of the participants going through the different statements and discussing them with the facilitator to ensure the correct understanding.

The character descriptions provided to the participants must be assumed as having a big effect on how the participant perceive the character. This includes the graphical presentation of the character. The study included a description of a consultant that had many years of experience. Being older of age, there is a possibility that the participants would find the character fairly conservative and base their answers on this. The graphical presentation of the consultant tried to combat this effect by including a computer as part of the consultant character. However the study did not find if any effect was present. As for the PrescriptAid character, it was mentioned by a participant whether the system also found information from the national interaction database (Interaktionsdatabasen.dk). This was not included in the character description, but it was hinted that the system knew about medical interactions by the graphical presentation as well as the advice provided by the system. This information could have been added to the description to create a more complete representation of the system.

The answers on the exit interviews, seen in [section L.3](#), also showed seven reports of concentration issues from the participants side. The issues could stem from the experiment being different in duration which is based off of the participants' answer speed. Possible solutions could have been more breaks and shorter active periods, this would however prolong the total amount of time the participant had to use on the experiment itself. Another idea could be to change the setup, and include some sort of gamification aspect to the experiment to motivate the participant and thereby minimising the effect of concentration issues. As mentioned in [subsection 5.2.1](#), one of the main problems by creating a paired comparison study was the number of participants needed. The study showed minimal problems concerning stochastic transitivity violations compared to the number of triads, and also fitted both a BTL model and a preference tree. However, making conclusions on these results should be made carefully the requirement of $6 \cdot n$ subjects, where n is the amount of stimuli, in this case 20. The experimental study should have included 120 participants. This study included 33 subjects, and since 5 were excluded the total number of subjects for used in the data analysis were 28 subjects.

Exploratory and experimental mixed method

The overall project included two different studies (exploratory and experimental), and these could have been done separately with a possibly larger scope for both. The exploratory study was crucial in creating the context in which the experimental study was carried out as well as the development and design of the decision aids and the description of the consultant. It is likely that similar changes to the proposed decision aid designs could have been found using more traditional software development techniques such as contextual design ([Holtzblatt & Beyer, 2016](#)) with iterative design sprints and validation through prototyping. It is however

found that indirectly measuring healthcare professionals' perceptions of decision aids through a pairwise comparison experiment did yield useful data that enables development of decision aids grounded in the user data.

6.6 Future work

As mentioned, the number of participants is low (28) compared to the suggested amount of 6 participants per stimuli (120). for using this kind of method. One of the possible solutions to the problem of recruiting participants, could be to create an online version of the experimental study. This would sacrifice some control of some otherwise controllable variables, such as making sure subjects take breaks, and subjects have read the character descriptions before initialising the study. This could extend the range of recruitment and also potentially increase the number of participants by allowing them to run the experiment at home or wherever they please.

This thesis focused on experts trust in decision aids and other experts, however for future work different subject groups could be of interest for comparing the results found in this thesis. These include how laypeople perceive experts vs decision aids and if these results mirror the results found in Dietvorst et al. (2015, 2016); Logg et al. (2019); Prahl and Van Swol (2017) that show a general trust towards decision aids. As described two of the four pillars of the danish hospitals are the administration and IT departments, these look to decision aids as tools for improving efficiency, cheaper healthcare and improving patient security. Uncovering how they perceive decision aids and experts could prove useful when discussing possible design solutions for improving the decision aids. Another interesting aspect to change in future work would be the gender of the human consultant. The study could be recreated with the same parameters but instead using a female human consultant to compare the results from the study. This could possibly shed some light on gender differences in the medical world and how these relate to decision aids.

As discussed, the studies that show that expert professionals distrust any advice, from decision aids or other people are based on *Weight of Advice* studies but these studies do not involve healthcare professionals. The exploratory part of this thesis found that healthcare professionals were quite willing to take advice from fellow colleagues, however this thesis did not shed any light on whether the advice is followed. Hence an experimental study using a *Weight of Advice* setup could be done to uncover whether this is actually true. As discussed earlier, different design changes were suggested that can possibly enhance the trust in decision aids. These different design changes could be the focus of further research, by trying to test whether they actually have an effect on the perception of the same statement for the decision aid. A study like this would follow the same setup but instead use a decision aid consisting of the changes by based on the changes described above. Through this thesis a total of 23 different statements were presented. These could be the focus of further research as the different nuances could provide an effect on the results reported. As mentioned, the economical incentive can possibly be perceived in different ways, and should be researched further before conclusions can be made.

6.7 Conclusion

To explore healthcare professionals' trust in decision aids and other healthcare professionals as well as which factors affect these perceptions, an experimental pairwise comparison study with 28 healthcare professionals and medicine students is conducted. The study is based on 10 perceptual statements about two adviser characters (a consultant and a decision aid) that were found, developed, and described based on semi-structured and contextual interviews with healthcare professionals and professors specialising in the Danish healthcare sector.

The experimental study shows that healthcare professionals perceive statements to be true for the decision aid to a larger extent than the consultant for 3 statements. Therefore, the decision aid is perceived to give advice based on scientific evidence, able to adapt to new knowledge, and have an economical incentive to a larger extent than the consultant. Healthcare professionals have likewise perceived the statements to be true for the consultant to a larger extent than the decision aid for 5 statements. Therefore, the consultant is perceived to give advice based on factors not scientifically provable, based on a complete picture of the patient, is responsible, takes changes to the patient's needs into account, and is able to explain his reasoning to a larger extent than the decision aid. The statements are not likely to be significantly different in 2 statements. Therefore, the decision aid and consultant is, by healthcare professionals, per-

ceived to be able to enhance the expertise of others and able to keep information confidential to the same extent.

Based on the combined results of the exploratory and experimental studies it is concluded that healthcare professionals are not averse to advice from decision aids, according to the three factors perceived as more true for the decision aid. Furthermore, it is found that the decision aid can be improved on the remaining 8 factors to perform as well as or better than the consultant. Changes to the design of the decision aid are suggested. These include putting larger emphasis on the amount of data taken into account by the decision aid, better visualisations and simulations of advice and consequences, providing the most relevant data at all times, highlighting the possibility of including devices and data not commonly thought of as collectable, and visualising an actual avatar of the patient as the centre of the decision aid to enhance the perception of taking the needs of the patient into account, basing advice on a complete picture of the patient, and better explaining the reasoning.

The mixed method approach to explore people's perception of algorithms and algorithmic advice in the form of decision aids is found to provide useful results. The results provide insights into factors that affect healthcare professionals' perception of advice from different advisers. The results can be turned into useful improvements and considerations for the design process of decision aids and algorithmic advice.

References

- Aalborg Havn. (2018). *Limfjorden bliver testområde for danmarks første førerløse havnebus.* [https://aalborghavn.dk/nyheder-\(1\)/limfjorden-bliver-testomraade-for-danmarks-foerste.aspx](https://aalborghavn.dk/nyheder-(1)/limfjorden-bliver-testomraade-for-danmarks-foerste.aspx).
- Andreassen, S., Zalounina Falborg, A., & Nielsen, A. D. (2019). *Treat.* <https://vbn.aau.dk/da/projects/treat-et-beslutningsst%C3%B8ttesystem-til-st%C3%B8tte-for-r%C3%A5dgivning-ved-an>.
- Bentley, E. (2019). *oTranscribe.* "<https://otranscribe.com>". MuckRock Foundation.
- Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational behavior and human decision processes*, 101(2), 127–151.
- Brinkmann, S., & Tanggaard, L. (2010). *Kvalitative metoder: en grundbog.* Hans Reitzels Forlag.
- Choisel, S., & Wickelmaier, F. (2005). Ratio-scaling of listener preference of multi-channel reproduced sound. *FORTSCHRITTE DER AKUSTIK*, 31(2), 793.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American psychologist*, 34(7), 571.
- Department of Health Science and Technology. (2019). *Glucosafe.* <https://www.hst.aau.dk/research-groups-centres/center-model-based-medical-decision-support/glucosafe/>.
- Diab, D. L., Pui, S.-Y., Yankelevich, M., & Highhouse, S. (2011). Lay perceptions of selection decision aids in us and non-us samples. *International Journal of Selection and Assessment*, 19(2), 209–216.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2016). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3), 1155–1170.
- Eastwood, J., Snook, B., & Luther, K. (2012). What people want from their professionals: Attitudes toward decision-making strategies. *Journal of Behavioral Decision Making*, 25(5), 458–468.
- Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational behavior and human performance*, 7(1), 86–106.
- Einhorn, H. J. (1986). Accepting error to make less error. *Journal of personality assessment*, 50(3), 387–395.
- Ellermeier, W., Mader, M., & Daniel, P. (2004). Scaling the unpleasantness of sounds according to the btl model: Ratio-scale representation and psychoacoustical analysis. *Acta Acustica united with Acustica*, 90(1), 101–107.

- Eriksen, H., & Ulrichsen, H. (1991). Tre kulturer i hospitalssektoren. *København: Nyt Nordisk Forlag.*
- Ernst, E., Merola, R., & Samaan, D. (2018). The economics of artificial intelligence: Implications for the future of work. *ILO Future of Work Research Paper Series No, 5.*
- Fry, B., & Reas, C. (2019). *Processing*. <https://processing.org/download/>.
- Gigerenzer, G. (2015). *Simply rational: Decision making in the real world*. Evolution and Cognition.
- Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: from innumeracy to insight. *Bmj*, 327(7417), 741–744.
- Goldman, L., Caldera, D. L., Nussbaum, S. R., Southwick, F. S., Krogstad, D., Murray, B., ... others (1977). Multifactorial index of cardiac risk in noncardiac surgical procedures. *New England Journal of Medicine*, 297(16), 845–850.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *science*, 315(5812), 619–619.
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, public policy, and law*, 2(2), 293.
- Haenssle, H., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., ... others (2018). Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology*, 29(8), 1836–1842.
- Harari, Y. N. (2016). *Homo deus: A brief history of tomorrow*. Random House.
- Harell, E. (2016). Managers shouldn't fear algorithm-based decision making. <https://hbr.org/2016/09/managers-shouldnt-fear-algorithm-based-decision-making>.
- Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, 1(3), 333–342.
- Holtzblatt, K., & Beyer, H. (2016). *Contextual design: Design for life*. Morgan Kaufmann.
- Hoy, M. B. (2018). Alexa, siri, cortana, and more: An introduction to voice assistants. *Medical reference services quarterly*, 37(1), 81–88.
- Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., & Nass, C. (2015). Why did my car just do that? explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 9(4), 269–275.
- Krantz, D., Luce, D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement, vol. i: Additive and polynomial representations*.
- Logg, J. M., Minson, J. A., & Moore, D. A. (2018). Do people trust algorithms more than companies realize? <https://hbr.org/2018/10/do-people-trust-algorithms-more-than-companies-realize>.
- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103.

- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press.
- Mendoza, M. (2014). Google develops contact lens glucose monitor. <https://news.yahoo.com/google-develops-contact-lens-glucose-monitor-000147894.html?guccounter=1>.
- Mermaid Care. (2019). BEACON Caresystem. "<http://beaconcaresystem.com/>". Author.
- Møller, A. (2016). *Evaluation of 3d positioned sound in multimodal scenarios* (Doctoral dissertation). (PhD Supervisor: Prof. Dorte Hammershøi, Aalborg University Co-PhD Supervisor: Assoc. Prof. Flemming Christensen, Aalborg University) doi: 10.5278/vbn.phd.engsci.00146
- Norman, C. D., & Skinner, H. A. (2006). ehealth literacy: essential skills for consumer health in a networked world. *Journal of medical Internet research*, 8(2), e9.
- Oberfeld, D., Hecht, H., Allendorf, U., & Wickelmaier, F. (2009). Ambient lighting modifies the flavor of wine. *Journal of Sensory Studies*, 24(6), 797–832.
- Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, 22(4), 390–409.
- Oxford Dictionaries. (2019). Big data. https://en.oxforddictionaries.com/definition/big_data.
- Park, J., Kim, J., Kim, S.-Y., Cheong, W. H., Jang, J., Park, Y.-G., ... others (2018). Soft, smart contact lenses with integrations of wireless circuits, glucose sensors, and displays. *Science advances*, 4(1), eaap9841.
- Petersen, L. S., & Bertelsen, P. (2012). A culture and power perspective on the management of health information technology in hospitals. In *Studies in health technology and informatics* (pp. 871–875). IOS Press.
- Petersen, L. S., & Bertelsen, P. S. (2017). *Equality challenges in the use of ehealth*.
- Petiot, J.-F., & Groonet, S. (2002). A multidimensional scaling approach for product design and preference modeling. In *Ieee international conference on systems, man and cybernetics* (Vol. 4, pp. 6–pp).
- Prahl, A., & Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, 36(6), 691–702.
- Promberger, M., & Baron, J. (2006). Do patients trust computers? *Journal of Behavioral Decision Making*, 19(5), 455–468.
- Rees, S. E. (2011). The intelligent ventilator (invent) project: The role of mathematical models in translating physiological knowledge into clinical practice. *Computer methods and programs in biomedicine*, 104, S1–S29.
- RStudio. (2019). Rstudio. <https://www.rstudio.com/products/rstudio/download/>.
- Rumelhart, D. L., & Greeno, J. G. (1971). Similarity between stimuli: An experimental test of the luce and restle choice models. *Journal of Mathematical Psychology*, 8(3), 370–381.
- Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., & Medow, M. A. (2013). Why do patients derogate physicians who use a computer-based diagnostic support system? *Medical Decision Making*, 33(1), 108–118.

- Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2017). Psychological roadblocks to the adoption of self-driving vehicles. *Nature Human Behaviour*, 1(10), 694.
- Showell, C., & Turner, P. (2013). Personal health records are designed for people like us. *MedInfo*, 1037.
- Sieck, W. R., & Arkes, H. R. (2005). The recalcitrance of overconfidence and its contribution to decision aid neglect. *Journal of Behavioral Decision Making*, 18(1), 29–53.
- Sundheds- og Ældreministeriet. (2018a, October). *Høreområdet i fremtiden: En styrket indsats for borgere med høretab*.
- Sundheds- og Ældreministeriet. (2018b). *Ét sikkert og sammenhængende sundheds-snetværk for alle - strategi for digital sundhed 2018-2022*. Sundheds- og Ældreministeriet, Finansministeriet, Danske Regioner, KL.
- Tornbjerg, K., & Bertelsen, P. (2014). *Undersøgelse af borgernes anvendelse af sundheds-it i 2013:-en udforskning af danskernes kendskab, holdninger og forhold til it, til gavn for eget helbred*. Dansk Center for Sundhedsinformatik.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, 79(4), 281.
- Waymo. (2019). Waymo. <https://waymo.com/>.
- Wickelmaier, F. (2019). *eba: Elimination-by-Aspects Models*. <https://cran.r-project.org/web/packages/eba/>.
- Wickelmaier, F., & Schmid, C. (2004). A Matlab function to estimate choice model parameters from paired-comparison data. *Behavior Research Methods, Instruments, and Computers*, 36(1), 29–40. doi: 10.3758/BF03195547
- Wolf, J. R. (2014). Do IT students prefer doctors who use IT? *Computers in Human Behavior*, 35, 287–294.
- Zimmer, K., & Ellermeier, W. (2003). Deriving ratio-scale measures of sound quality from preference judgments. *Noise Control Engineering Journal*, 51(4), 210–215.
- Zimmer, K., Ellermeier, W., & Schmid, C. (2004). Using probabilistic choice models to investigate auditory unpleasantness. *Acta acustica united with acustica*, 90(6), 1019–1028.

Appendix

Interview I

A

This appendix contains the interview guide, the transcription and the analysis for interview I.

A.1 Interview Guide

This section contains the interview guide for the first semi-structured expert interview conducted during this study. The interview guide is seen in [Table A.1](#).

Table A.1: Interview guide I

Introduktion: Interviewet omhandler din viden om sundhedsdomænet og har til formål at kortlægge domænet samt de karakterer – eksempelvis eksperter, lægmænd eller algoritmer – der er relevante i forbindelse med folks opfattelse af råd og vægtning af disse.

Anonymitet: Dit navn anonymiseres, hvis dette ønskes.

Båndoptagelse: Interviewet optages på bånd og transskriberes.

Tid: Interviewet vil tage maks en time.

Forskningsspørgsmål

Hvem er interviewpersonen og hvad er interviewpersonens kvalifikationer til at være ekspert i området? Har interviewpersonen et overblik over sundhedsområdet? Er interviewpersonen en del af sundhedsområdet, eksempelvis en del af hierarkiet?

Hvilke *rådgivere* – fysiske attributter – er relevante i forbindelse med folks opfattelse og vægtning af råd? Hvordan kan disse karakterer beskrives?

Stemmer ansvarsfordelingen af forskellige rådgivere overens med patientens opfattelse og vægtning af råd? Hvem står for at tage beslutninger i forløbet?

Hvordan beskrives rådgiveres beslutningstagen for patienter? Beskrives forskellige rådgiveres beslutningstagen forskelligt? Hvordan opfatter patienten forskellige rådgiveres beslutninger? Hvad er vigtige faktorer i forhold til at føle tillid til beslutninger/råd?

Hvad er sundhedspersonales indtryk af algoritmer og software? Ses software som hjælpemidler til beslutningstagen? Er der tiltro til at software kan give retvisende råd?

Interviewspørgsmål

- Hvad er din baggrund/uddannelse?
- Hvad er dit forskningsområde?
- Hvor lang tid har du arbejdet med/i sundhedsområdet?
- Hvorfor startede du i sin tid med at forske i dette område?
- Har du arbejdet med forskellige personale i sundhedssektoren (Læger, sygeplejesker, fysoiterapeuter, diætister, etc.)?

- Hvilke roller findes der i sundhedssektoren, i forhold til behandling af patienter?
- Kan du beskrive noget der karakteriserer de forskellige roller? (Hvad går de ud på/Hvad laver de?)
 - Hvad karakteriserer en læge, en sygeplejerske, en diætist, etc.?
- Hvordan er hierarkiet af rollerne indenfor patientbehandling?
- Er der nogle fastsatte rammer for de forskellige roller?
 - Er der nogle definerede arbejdsopgaver for de forskellige roller?
 - Er der nogle uskrevne rollefordelinger og arbejdsområder?
- Fungerer alle roller som rådgivere for patienten?
- Er det dit indtryk at patienter opfatter og vægter råd fra forskellige roller forskelligt?
 - Kan du beskrive en situation, hvor dette er tilfældet? (eg. læge og diætist, eller læge og sygeplejerske)
 - Kan dette ske at være en effekt af det hierarki der er præsenteret for patienten?
- Hvem står for beslutningstagningen?
 - Kan du nævne nogle situationer, der beskriver eksempler på beslutningstagningen?
 - Hvem står som ansvarlig for patienten? (Eg. læge, sygeplejerske, diætist, etc.)
 - Er patienten en aktiv del af beslutningsprocessen?
 - Benyttes der teknologi – *algoritmer* – til beslutningsprocessen?
- Kan du komme i tanke om, forskellige opfattelser patienter kan have, som påvirker patientens tillid til råd og beslutninger?
 - Kan du eksempelvis beskrive en situation, hvor autoritet påvirker patientens tillid?
 - Kan der være forskel på tillid til forskellige eksperter baseret på deres autoritet?
- Gøres der noget for at beskrive beslutningsprocessen til patienten?
 - Beskrives forskellige rollers beslutningstagen forskelligt? (Eg. en diætist og en læge)
- Hvordan opfatter patienter forskellige rollers beslutninger/råd?
 - Kan du komme i tanke om nogle faktorer, der påvirker patientens følelse af tillid til beslutninger og råd?

- Hvad er sundhedspersonales indtryk af algoritmer og software?
 - Kan du nævne en situation, hvor sundhedspersonalets holdninger kommer til udtryk?
 - Ses software som hjælpemidler til beslutningstagen?
- Er der tiltro til at software kan give retvisende råd?
 - Kan du beskrive, hvilke faktorer, der spiller ind i forhold til om sundhedspersonale har tiltro til software?

A.2 Transcription

This appendix contains the transcription of interview I. The transcription was made using the oTranscribe website ([Bentley, 2019](#)).

Transcription

00:43 Mikkel: Bare lige for at starte fra starten her, hvad er din baggrund / uddannelse?

00:55 IP: Altså jeg er uddannet på RUC, noget der hedder teknologi og samfundsplanlægger tilbage i 87, og så har jeg arbejdet med infrastruktur lige siden, men på forskellige måder, men de sidste 20 år har jeg arbejdet med sundhedsinformatik. Altså IT inden for sundhedssektoren.

01:10 Mikkel: Ja. Har du arbejdet med forskellig personale inden for sundhedssektoren i den her tid så?

01:24 IP: Altså jeg har lavet, jeg har arbejdet rigtig meget med lægesekretærer, og har lavet nogle studier med lægesekretærer det ved jeg ikke om i har googlet på VBN?

01:35 Mikkel: ja.

01:35 IP: Men også lavet mere sådan, være involveret i sådan nogle evalueringer af EPJ-systemer og pilottest af EPJ-systemer og sådan noget, men jeg har arbejdet meget med metoder til at få, og finde ud af hvad er det folk gør når de gør deres arbejde.

01:54 Mikkel: Ja.

01:54 IP: Og hvordan kan man bedst få det frem, øhhh, og meget de sidste ti år her eller frem handler det om design, altså, hvordan kan man simpelthen inddrage brugeren i design af IT.

02:04 Mikkel: Ja. Øhhmm, så lidt videre til det her med roller i sundhedssektoren, hvad for nogle roller findes der? I forhold til behandling af patienter.

02:04 IP: Det er meget bredt altså, æhmm, der findes en bog der hedder 'tre kulturer i sundhedssektoren' som er skrevet, det er sådan en projektledelsesbog, jeg har den oppe på mit kontor i kan se den, som er skrevet for måske tyve år siden eller sådan noget, hvor man snakker om der er lægerne, der er sygeplejerskene og så er der djøfferne. Men siden da er der jo kommet flere kulturer til og det er jo IT-ekspertene. Så jeg har så sammen med en kollega skrevet en artikel, hvor vi så snakker om de fire kulturer i sundhedssektoren ikke, fordi at IT-afdelingen får jo mere og mere magt. Men der er jo også sket det at fra at IT-afdelingen var på sygehøsten, så er den jo nu ude i regionen, så nu er det jo egentlig fjernet fra klinikken, nu er det jo blevet så centralt at de sidder et andet sted ikke. Så har man jo selvfolgtlig noget daglig drift og nogle superbrugere, ligesom vi har et kontor hernede ikke men i virkeligheden sidder de jo ude på campus.

03:20 Mikkel: Ja.

03:20 IP: Men så sidder der bare så to-tre forkølede personale hernede ikke, som så skal servicer os allesammen. Men egentlig decoupled fra selve hovedafdelingen, sådan er det også for sygehusene. Ja, og men det var sådan mere forskellige kulturer. Fagkulturer, og dit spørgsmål gik mere på?

03:44 Mikkel: Ja nu nævnte du læger og sygeplejerske for eksempel, er der andre sådan.

03:49 IP: Ja så er der lægesekretærer.

03:49 Mikkel: Er det også øhh.

03:52 IP: Altså der vil jeg så sige, hvad er det for nogle folk fra sygehuse der arbejder med data, for det er jo det der er interessant når vi snakker om algoritmer ikke, er det ikke data? Dataarbejde der er interessant eller hvad? Altså man skal jo ligesom sidde foran en computer.

04:11 Mikkel: Ja, ja.

04:13 IP: Altså det er jo ikke nogen der står og skærer maven op på nogen som så, ja altså jo selvfolgeligt, hvis det er en robot. Men det handler vel om noget dataarbejde ikke?

04:20 Mikkel: Ja.

04:20 IP: I snakker om det som noget hierarkier eller hvad?

04:28 Mikkel: Ja eller det kommer vi lidt til senere, men hvilke roller er der, hvordan kan man sådan karakterisere de forskellige roller, hvad karakteriserer lægen i modsætning til sygeplejersken og i modsætning til lægesekretæren. I forhold til hvad de laver. Og hvad deres roller er og hvad deres opgaver er.

04:45 IP: Det er jo meget meget bredt, men altså man kan sige at så er lægerne jo de videnskabeligt funderede sådan en videnskabelig fundert personalegruppe ikke, altså som hævder at de har evidens for hvad de gør og de gør kun tingene hvis der er evidens for det, ikke, det er ligesom den positivistiske side, hvor sygeplejerskerne er mere hermaneutiske i deres tilgang, altså de er mere sådan det hele menneske og omsorg og sørge for at levere en professionel pleje der får såret til at hele som lægen har lavet på patienten ved at skære på ikke. Øhhmm. Så man kan sige der er, øhhm. Altså man kan sige de har nogle forskellige baggrunde for at

gøre hvad de gør, men det er jo klart lægen der har magt på sygehusene. Altså det er jo dem der har, øhhmm, altså det er jo ligesom dem der bestemmer, fordi det er dem der har ansvaret. Det der så er lidt sjovt det er at hvis man har læger og sygeplejersker og lægesekretærer, så er der meget større alliance mellem lægerne og lægesekretærerne fordi lægesekretærerne traditionelt har været dem der har hjulpet lægerne med at dokumentere. Og jeg ved ikke om i har fulgt med, men ellers så kan i søge det frem, alt den ballade der har været omkring sundhedsplatformen på Sjælland, har jo handlet om at man har lavet skodder mellem hvem der kan se hvad i den elektroniske patientjournal, sådan så lægesekretærerne ikke længere kan hjælpe lægerne. Og det har man fordi man anvender et der er gammelt og i sin rod er bygget op efter det amerikanske sundhedssystem. Hvor man for eksempel ikke har lægesekretærer men man har ti forskellige andre funktioner og personalegrupper som laver det samme som vi har lægesekretærer der laver i dag. Og derfor så fungerer det ikke i en dansk kontekst, så det er et glimrende eksempel på teknologioverførsel der ikke virker. Så har man så lavet en masse modifikationer og det er ikke så slemt som det var i starten og har fundet ud af sådan nogle workarounds, hvordan man ligesom kommer omkring og sekretæren så kan gå ind og hjælpe lægerne, det de var, fordi lægerne skulle jo selv skrive, sekretærerne har jo egentlig været kvalitetstjekkerne på det lægerne lavede, fordi vi har et system i Danmark, hvor man er lægeseretær på en afdeling, dvs. hvis man er på en børneafdeling, så har man måske været der i ti år og så skriver man for den læge der har med børn, der har med diabetes at gøre. Så når hun har et ambulatorie, hvor alle de der diabetesbør bliver set, så sidder der en sekretær og skriver for det. Og det vil sige at sekretæren har en fagfaglig forståelse. Hun er ikke bare en HK'er der sidder med dutter i ørene og skriver, hun har faktisk en forståelse for området, og det gør, at når lægen så bliver forstyrret i sin diktering og kommer til at sige noget forkert, eller der mangler et ord, fordi man har spolet for langt eller sådan noget, så kan sekretæren fange det og gå hen og sige, 'hey hvaøhh', 'gud ja der skulle stå ikke' eller hvad det nu måtte være ikke. Så på den måde har der været et meget tæt samarbejde som man jo så er ved at splitte, altså, og ødelægge. Fordi at dem der sidder og laver systemerne ikke ved hvad der foregår på sygehusene, det er måske sådan nogle som jer som kommer ud og forstå hvad der foregår ikke. Øhhm, fordi man ikke går ud og undersøger praksis men gør det på fornøyelsens arken, fordi i USA er det så godt osv. ikke? Så men man kan sige at der har ligesom været et tæt samarbejde og et tæt kvalitetssikringssamarbejde mellem læge og lægesekretær. Og det er der så andre steder hvor det stadigvæk findes, men IT kommer selvfølgelig ind og rykker rundt på mange af de ting ikke, hvor sygeplejerskerne ligesom har deres eget og så kan der selvfølgelig være noget jobglidning ind imellem og så videre men de kører ligesom, de har aldrig haft sekretærer til at hjælpe sig med deres arbejde, så, men jeg ved ikke om det er svar på dit spørgsmål, for jeg synes at det er svært det med hierarkier.

09:24 Mikkel ja det er det.

09:24 IP: Ja altså man har jo på en sygehusafdeling altid en overlæge, eller professor eller hvad det nu er. Som afdelingsleder og en oversygeplejerske som afdelingsleder der er de to ledere på sådan en afdeling, der ligesom afspejler. Og noget af det som har været rigtig svært, er at lægesekretæren er ikke anerkendt som en del af afdelingsledelsen selvom at det egentligt er dem der styrer rigtig meget af det sundhedsadministrative og at det er blevet meget mere så har man stadigvæk blandt eksperterne, øhhm, hvad hedder det, administratorerne, embedsfolket, har man stadigvæk en forestilling af at de sidder med dutter i ørerne og skriver alt muligt blandet i sådan en kælder ikke. Så. Så der er sådan et hierarki men altså i kunne jo prøve at kigge den der bog, øhhm, om de der forskellige kulturer for det beskriver jo meget godt. Og man kan sige der hvor djøfferne så kommer ind det er jo hele altså deres indflydelse, lægerne hader dem jo som pesten, fordi det er dem der kommer og bringer embedsværkets, ministeriets, øhh, ligesom skal, de skal piske lægerne på plads til at få gjort og så videre ikke. Og det er jo blevet mere og mere fordi at man prøver at detailstyre sundhedsvæsnet fra statens sida af og new public management og jeg skal komme efter jer ikke. Det er man så ved at være færdig med på sygehusene nu, nu snakker man værdibaseret ledelse og sådan noget ikke, mens man her kører på 200 kilometer i timen ikke og decentraliserer alting ned til den mindste forskningsgruppe ikke, så vi dårligt kan, altså, alle skal, det er fuldstændig ude af trit, men det har man ikke opdaget på universitet endnu, det skal nok komme om en ti år. Men altså det er bare sindsygt at være vidne til.

11:17 Mikkel: Du snakker lidt om det her med at lægen skærer og sygeplejen passer det sår, er der andre sådan fastsatte rolleopgavefordelinger for de her forskellige roller karakterer, hvad skal man sige? Veldefinerede arbejdsopgaver, hvis man kan sige det sådan.

11:34 IP: Det er rigtig rigtig meget kontekstbestemt, det kan jeg simpelthen ikke sidde og sige noget om, for det vil afhænge af, hvad det er for en afdeling, altså det må i næsten, så skal i sige, hvis de er på skadestuen, ja så snakker vi om skadestuen, eller hvis det er på en medicinsk afdeling så snakker vi medicinsk, hvis det er på en kirurgisk, altså det er helt forskelligt, øhh, hvad man laver og hvem der gør hvad, altså man kan sige at generelt så er det mere kompliceret

på medicinske afdelinger fordi det handler om sygdomme inde i kroppen, mens på kirurgiske der har man jo fundet ud af hvad sygdommen er så skærer man det væk eller syer det sammen eller lapper det, og der er det på en måde mere enkelt men så er der jo alt muligt andet der kan være, altså en hjernekirurgisk er jo ikke enkelt, så det øhh, det er rigtig svært at sige men man generelt, så, øhh, så er det mere kompliceret, øhh, der er, øhh, de medicinske læger bruger meget tid på at finde ud af, hvad patienten fejler. Og ikke så meget tid på behandlingen, mens de kirurgiske, de bruger meget lidt tid på at finde ud af, hvad patienten fejler, for det er måske endda nogle andre der har fundet ud af, og så bruger de meget tid på selve behandlingen og operere og komme sig og altså, men det har jo ikke noget med hierarkier at gøre. Det er bare noget med arbejdsspraksis der er forskellige ikke.

13:14 Mikkel: Ja, ja.

13:14 IP: Og man kan sige sygeplejerskernes arbejde er jo også forskelligt, altså sygeplejersker er jo også specialiserede og i dag kan man jo læse kandidat i sygepleje enten inden for noget administrativt eller inden for, altså man kan jo være en specialiseret hjertesygeplejerske der ved alt om hjertet og så er man jo også næsten på niveau med, altså, med læge og så kan man være en der er på en almen medicinsk afdeling som ligesom mere tager sig af fordi folk har feber og altså hvad ved jeg, så det kommer også an på hvad de arbejder med. Det er simpelthen bare ikke nemt at sige noget om.

13:46 Morten: Hvad så med sådan noget som diætister og fysioterapeuter og sådan noget, hvordan passer de ind i, øhh.

13:53 IP: Altså der er mange forskellige, øhh, vi lavede engang sådan en, med nogle studerende, med hvad findes der af personaler på sygehuse, der har vi faktisk sådan en hel tavle med billeder, det kan jeg måske prøve at sende til jer hvis det er, men men øhh, der er jo alt muligt forskelligt, men i hierarkiet, altså vi har et meget flat hierarki i Danmark i forhold til USA ikke, hvor det der med at stille den enkelte til ansvar er ikke noget som bliver praktiseret på samme måde og jeg ved ikke om i kan huske og hørte efter dengang der var noget der hed Svendborg-sagen her for nogle år siden. Med en ung læge nede på Svendborg sygehus som havde været i vagt i 24 timer og havde sagt til sygeplejerske at de skulle tjekke en patient der var blevet indlagt på akut-medicinsk afdeling at de skulle tjekke hans blodsukker, og det havde sygeplejersken ikke gjort og det viste sig at han var diabetiker og han var bevidstløs, eller hvad han nu var, i hvert fald ikke ved sin, kunne ikke selv gøre opmærksom på det, og han endte med at dø fordi de ikke havde opdaget at han havde diabetes og han var gået sukkekold eller fanden, eller for meget, jeg ved det ikke.

15:06 Mikkel: Ja.

15:06 IP: Og så blev den nye styrelse for patientsikkerhed anklagede så lægen for at have slået ham ihjel, og det var helt nyt i Danmark. Det plejer man ikke at gøre, fordi vi har et meget tillidsfuldt, øhh, vi har faktisk brugt rigtig mange år på at bygge en kultur, hvor man indrapporterer utilsigtede hændelser, og så gør det anonymt eller ikke anonymt, fordi vi gerne vi fange fejlene, hellere det end vi vil blame nogen ikke. Så det case har, og det endte jo så med efter et år, og så kom den der, 'det ku have været mig' kampagne med alle lægerne der skrev osv. altså der var sådan en kæmpe gang i faget ikke, og det endte med at hun blev frikendt. Men altså hun fik jo smadret sit liv af hele det der ikke, så hun skal ikke have med patienter at gøre, hun er blevet forsøker i stedet for, ikke. Så, øhhrm, men hvad ville jeg sige med det, hvorfor sagde jeg det?

16:06 IP: At der sker en ændring måske i den kultur, hvormed man praktiserer sit fag ikke. Ja. Men det med andre professioner, altså der kan man jo, altså der er ikke noget hierarki, altså fysserne har jo deres egen afdeling, der hvor der har slået nogle stag, har været slået nogle slag, det har været omkring for eksempel designet af den elektroniske patientjournal, hvor meget skulle fysioterapeuter og ergoterapeuter have lov til at se og hvem skulle have lov til at se hvad de skriver og altså, der har så foregået nogle forhandlinger. Og tilretninger osv. ikke. Fordi ikke. Men der har det jo mest været tavshedspligten man har fået efter, fordi de ikke havde samme form for autorisation eller så ved jeg ikke om det er blevet ændret. Så hele det der med hvad der findes af folk af sygehuse, så er der jo laboranter eller bioanalytikere som de hedder nu og der er jo massere forskellige, portører og, ja.

17:12 Mikkel: Ja.

17:12 IP: Men det er jo sådan i dag, at man kan jo ikke gå ind og slå op på en patient. Altså hvis du ikke har behandlingsretten, så hvis du slår op på, fordi du ser din nabo kommer gravid ind, og du så efter otte timer vil se fordi du er på en anden afdeling, vil se fordi hun har født, om hun har fået en pige eller en dreng, og du så slår op på journalen på hende, så bliver du fyret lige med det samme. 'fløjt'. Det må du ikke, fordi du har ikke noget med hende at gøre, og den er helt klar og hård, og det er jo det samme hvis jeg har en veninde der er læge og som jo nogengange har givet mig et eller andet, hvad ved jeg, så hvis hun logger på sundhed.dk for at se noget sammen med mig, så får jeg en mail om at hun har været inde og kigge på mig.

Fordi hun er ikke min behandler, men hvis min praktiserende lægehus eller nogle dermede gør det, så får jeg det ikke. Så det er det der er helt, sådan, og det er jo IT der gør det muligt, ikke.

18:12 Mikkel: Hvordan med de her forskellige roller, fungerer de allesammen som, skal vi kalde det rådgivere, for en patient, altså det her, hvis de skal give råd til hvordan patienten skal enten leve eller nogle forskellige ting.

18:28 IP: Altså læger og sygeplejerske går jo stuegang sammen, så på den måde så kommer de ind og står og snakker og så går de igen. Og så ser patienten ikke lægen mere, men patienten ser jo sygeplejersken igen så derfor så har man jo tættere kontakt så kan man spørge, hvad var det egentligt som hun sagde eller sygeplejersken kan 'nu skal du komme igen og forklare noget mere' ikke. Så de fungerer jo som et team.

18:58 Mikkel: Ja.

18:58 IP: Og det kan så organiseres på alle mulige forskellige måder. Altså nogle gange så er der en læge der går med dem allesammen og andre gange så har hver læge sine og sådan noget, det er også igen hvis man er på en kirurgisk afdeling så vil kirurgen typisk gå til de patienter som han eller hun har skåret i, så går de til deres egne med nogle sygeplejersker selvom patienterne så ligger spredt rundt på afdelingen så vil de gå med forskellige sygeplejersker der passer de patienter ikke.

19:22 Mikkel: Ja, ja.

19:24 Mikkel: Er det dit indtryk at patienter så opfatter og vægter rådene forskelligt fra, hvilken karakter, eller hvad siger man personale der så giver råd.

19:37 IP: Øhhm, altså jeg tror stadigvæk at at det lægen siger bliver vægtet højere, fordi lægen er specialist i det man fejler. Eller det forventer man i hvert fald at lægen er, ikke. Men det er ikke, altså den måde jeg lytter til hvad lægen siger er jo ikke med samme erbødighed som min mormor lyttede til hvad lægen sagde. Altså dem der har levet før penicilinen blev opfundet... *hehehe* ... havde jo ligesom en anden, øhhm holdning til, til lægen og det der jo sker i dag det er at rigtig mange patienter de spørger jo også google eller sundhed.dk.

20:12 Mikkel: Ja.

20:12 IP: Øhh, og det er med til at ændre, øhhh, relationen ikke, fordi man som patient kommer med spørgsmål. Og det er også sådan at sygehuse når de indkalder til nogle undersøgelser, så beder de sådanset patienterne om at skrive deres spørgsmål ned hjemmefra. Som en del af forberedelsen, der har de i hvert fald fået sådan et brev, ikke, med sådan en tjkliste og 'hvis der er nogle spørgsmål så skriv dem ned, så du ikke glemmer at stille dem' sådan at man inviterer til at folk de kommer med deres spørgsmål ikke. Så det mere er en gensidig dialog. Jeg tror godt at de fleste læger, de kender godt deres begrænsning, altså det der med at tro at læger de har styr på det hele, det har de ikke. Altså der er jo rigtig rigtig meget som der er ingen der ved, hvorfor er. Og selvfølgelig er der noget der er evidens for, men der er også rigtig meget som er sådan noget baseret på erfaringer, og vi ved det virker, vi ved ikke hvorfor det virker, men vi ved det virker, altså sådan er der jo rigtig meget behandling der er. Det ved vi almindelige mennesker bare ikke så meget om. Vel?

21:10 Mikkel: Nej.

21:08 Morten: Også for eksempel diætister, hvor de for eksempel skal ind og rådgive en patient, eller en læge skulle ind og rådgive, kunne der være noget der, der gør.

21:21 IP: At man lytter mere til lægen end til diætisten?

21:23 Morten: Ja, selvom måske diætisten har mere erfaring inden for det område eller et andet.

21:27 IP: Det ved jeg simpelthen ikke nok om. Altså, øhh, i det hele taget så, jeg tror, hvis jeg skal snakke ud fra mig selv, så øhh, så ville jeg nok finde ud af, hvad for en uddannelse den der diætist havde, fordi øhhm, hvis det var en sygeplejerske der havde specialiseret sig inden for ernæring ville jeg nok lytte mere til hende, end hvis det var en social og sundhedsassistent der havde. Men altså så ved jeg slet ikke om hun ville være ansat på sygehuset. Altså der er jo rigtig mange, af sådan nogle diætister, der kalder sig selv diætister som er sådan noget slankefis og hvad ved jeg ikke. Og dem ville jeg da ikke lytte til på samme måde som jeg ville lytte til en der var ansat på sygehuset, men jeg tror at hvis diætisten var ansat på sygehuset så ville jeg også lytte til hende for så er der nogle der har sagt god for hende og det betyder noget for mig. Altså at systemet har sagt god for hende ja.

22:26 Mikkel: Ja.

22:26 IP: Eller ham.

22:26 Mikkel: Hvad så hvis, hvis man står i sådan en situation, hvor de måske, en læge og en diætist, måske havde givet forskellige råd, eller forskellige retninger man skulle opføre sig i?

22:39 IP: Så ville jeg nok, øhhh, men det er ikke sikkert alle ville det, men så ville jeg nok lytte på diætisten, fordi diætisten er jo den der har studeret ernæring ned i mindste, hvor lægen, jeg ved ikke om man kan være specialiseret i ernæring som læge, det er jeg ikke sikker på at

man kan, men altså hvis det var, hvis jeg nu havde en tarmsygdom, øhhm, og læge ud fra de prøver der var taget kunne se at jeg ikke kan tåle gluten eller hvad det nu måtte være, og der så kom en diætist og sagde at det skulle jeg da ikke sådan spekulere over, så ville jeg jo stole på lægen, men altså man kobler jo ikke hjernen fra. Så jeg tror meget om det er bundet op på en bestemt sygdom eller at det er sådan et råd om, hvordan taber du dig bedst.

23:25 Mikkel: Ja.

23:25 IP: Altså der er meget forskel på om det, og hvordan man, tror jeg.

23:29 Mikkel: Ja.

23:33 Mikkel: Øhhmm.

23:35 Mikkel: I forhold til det her med, det har vi også lidt været inde på, men det her med beslutningstagnning og sådan noget, hvornår, øhhm, hvem der egentligt står for det i forskellige situationer.

23:45 IP: Altså det er jo altid lægerne i sidste ende der beslutter, hvad der skal ske, og har ansvaret. Og så er det jo på baggrund, altså kender i noget til arbejdsgangen? Altså de holder jo sådan nogle konference hver dag på sygehuset kalder de det og diskuterer de patienter der er kommet ind eller har.

24:03 Mikkel: Ja.

24:03 IP: Altså, så, så det gør, at de render rundt til hver deres patienter og så mødes de en gang om dagen, tit så starter de med at komme ned og have en morgenkonference. Men det er også afhængigt af hvad det er for en afdeling. Hvis nu det er en hjerteafdeling, hvad er der kommet ind i løbet af natten og hvad skal vi gøre med dem og så bliver de sendt rundt og så mødes man måske klokken 14 og holder sådan en eftermiddagskonference når man har fået alle data ind og hvad skal der så ske videre, men selvfølgelig, hvis det er en akutafdeling så har man det hele tiden ikke, men hvis det er, det er, øhh, men nu er der jo ikke noget der ikke er akut for ellers kan du jo ikke få lov at ligge på sygehuset nærmest, for man bliver jo sparket ud lige med det samme. Så, øhh, men hvis det nu er brækkede ben eller sådan noget, så holder man måske kun en om dagen, ikke, men hvis det er på intensiv, hvor at der er en læge og en sygeplejerske og alt muligt omkring hver patient så, øhh, hele tiden, ikke, så holder man oftere.

24:54 Mikkel: Ja.

24:54 IP: Jeg kan fortælle noget sjovt, øhh, som måske kan give lidt til det her med beslutninger, æhhm, på intensivafdeling, jeg lavede sådan nogle interviews med de der lægesekretærer, der har de, de har mangel på lægesekretærpersonale, øhhh, men der har man en praksis, hvor der er en journal, som bliver dikteret og skrevet og så har de en anden journal, de har sådan en dagsliste, de har sådan et stykke papir nede i lommen som også bliver skrevet af sekretærerne, men som ikke bliver offentliggjort på eJournalen. Eller bliver, men som er lægernes arbejdsredskab, og som de bruger og hvor de noterer alt muligt ned som de så snakker om på konferencen, og selvfølgelig har det noget at gøre med at når man ligger på intensiv så er man jo mega mega syg og det ændrer sig *knipser* fra minut til minut, hvordan folk har det, så de har brug for nogle opdaterede oplysninger, de kan ikke sidde og vente på at der er en sekretær der skriver og gør ved og væk ikke. Men alligevel, så skal de der lister være skrevet, altså de der sedler være skrevet inden klokken et eller andet. Øhh, de der interne lyster, og, så der kører de egentlig dobbeltjournalføring fordi lægerne har brug for den ene som deres arbejdsredskab og den anden som dokumentationen over for hele verden og omegnen ikke. Og om de så, hvad for en der betyder mest i forhold til beslutningerne, der tror jeg den der nede i lommen betyder mest, men altså det kan jeg jo selvfølgelig, det er jo bare et gæt, men sådan er der mange forskellige praksiser som knytter sig lige præcis til den funktion som man har.

26:29 Mikkel: Ja.

26:29 IP: Så det med at fjerne alt papir, altså 'nu skal vi gå papirløst' det er jo fuldstændig sindssygt, jeg ved ikke hvem fanden der finder på at det skulle være godt, altså hvad er der i vejen med papir? Papir kan nogle ting og så kan det digitale noget andet ikke.

26:43 Mikkel: Ja. Øhhm. I forhold til. Eller kan du nævne nogle situationer hvor sygeplejersken måske godt kan tage beslutningerne uden lægens direktioner...

26:58 IP: Ja, ja.

26:56 Mikkel: ...hvor også måske, hvor de kan have ansvaret for, hvis der så går noget galt. Er der nogle situationer, hvor det kan ske?

27:08 IP: Jeg sidder og tænker på sådan noget med at lægen, når lægen ordinerer medicin, og der er noget der hedder 'PN', og det betyder, og jeg ved ikke om det er latin, men det betyder 'efter behov' og så kan en patient som ligger og er blevet opereret, med sin hofte eller et eller andet, kan have panodil og ipren PN. Dvs. at hvis patienten har, altså de har fået ordinet til at skulle have det fire gange i døgnet, ikke, men hvis man har meget ondt så kan man få noget mere. Det ordinerer lægerne og så dispenserer, altså deler sygeplejerskerne det ud og der er det op til sygeplejersken at vurdere om patienten har så ondt at de skal have noget mere af det lægen siger at de godt må få ad libitum af, eller, altså, ikke?

27:52 Mikkel: Ja.

27:52 IP: Og hvis hun så giver for meget, så står hun jo til ansvar for der er jo, altså en eller anden tabel for hvor meget mennesker må have eller må få og der vil sygeplejersken så stå til ansvar. Altså der var også, kan jeg huske, der var også omkring, øhhm, for et par år tilbage, omkring medicinmodulet, så man havde heroppe på det tidspunkt, at der har der været meget kuk i med de der medicinmoduler om det er milliliter eller om det er mikrogram, eller det er, altså at fra at man går, altså at man skifter måleenhed, og så nogle gange kommer personalet til at tage forkert, og der var der en der blev slæt ihjel oppe i Dronninglund eller i Hjørring, fordi hun fik alt for meget morfin, fordi at sygeplejersken ikke havde forstået at det var en anden, altså og det er så også noget med nogle bestemte præparerater som man tidligere har givet i mikrogram og nu giver i milliliter eller et eller andet sådan noget, hvor man i en travl kan komme til at lave noget ikke. Men jeg ved ikke engang om hun, altså det skal jeg ikke kunne sige, men altså sådan noget kan der godt ske. Men altså lægerne uddelegerer jo langt hen ad vejen ansvaret til sygeplejerskerne. ... 29:11

29:15 IP: Og sådan er det.

29:19 IP: Altså der var også den der sag nede ved Nakskov med hende der der havde slæt alle mulige ihjel. Kan i huske det?

29:22 Mikkel: Ja det tror jeg godt jeg kan huske.

29:23 IP: Ja, her sidst, for et halvt år siden, altså en sygeplejerske som, øhh, at de skulle ikke ligge og lide der og så gjorde et eller andet, og det har hun jo adgang til i medicinrummet og der er noget svind og sådan ikke, men det der så er dokumenteret, altså det som lægen har dokumenteret det er jo det de skal gøre. Og så går man jo hen til den vagthavende læge, det er jo derfor der hele tiden er læger der er i vagt og bagvagt og sådan noget, for opstår der noget så ringer de jo og så kommer de ind og så ordinerer de noget andet, ikke.

29:53 Mikkel: Ja. Er der situationer, hvor patienten også kan være en del af den her beslutningsproces...

30:01 IP: Ja.

30:01 Mikkel: ...i forhold til behandlingen?

30:01 IP: Ja. Tit så vil man jo spørge patienten, hvad vedkommende selv synes og har erfaring med. Blandt andet fordi, at det jo ikke er alt man har på en formel, det er jo ikke matematik, altså, der er meget der også handler om patientens erfaring, 'har du nogen sinde fået det her medicin før?', spørger de, ikke. Og så kan man sige 'ja det har jeg for ti år siden', altså det kan de jo ikke se inde i det fælles medicinkort. Så. Øhhm. Ja det vil man gøre.

30:36 Mikkel: Og så over til det her med teknologien i forhold til alle de her medicineringer og ordineringer osv., hvor høj grad er der i den del, hjælpemidler til beslutningstagen?

30:51 IP: Altså, det er jo noget, hvis i læser, i kunne prøve at læse digitaliseringstrategien for sundhedsvæsnet som kom i 18', fordi der står faktisk, hvad man arbejder på, og en af de ting man arbejder på, det er at gøre det mere sikkert at, altså og, øhh, man vil gerne lave noget bedre beslutningsstøtte omkring præparerater som, hvad hedder sådan noget, som kontra, altså som ikke skal tages sammen, fordi det så enten ophæver virkningen eller gør noget, ikke. Det er en ting man arbejder på at lave noget beslutningsstøtte for det. Noget andet er at man godt kunne tænke sig at være bedre til, altså man vil godt have noget der hedder cave, som er til alle patienter, og det handler om om man har allergier overfor noget, og det kan være allergier overfor hø og hvad ved jeg, men det kan også være allergier overfor nogen, altså nogen kan jo ikke tåle penicilin for eksempel og nogle kan ikke tåle ditten og datten, ikke. Og i dag er det ikke sådan at det er automatisk, altså der er ikke nogen central. National database for det. Så det skal du ligesom oplyse, hver gang du kommer ind med det, altså selvfølgelig har man det i region Nord eller man har, øhh, det står i, jeg, øhh, jeg skal ikke kunne sige det, men jeg tror de spørger om det, hver gang. Men det er sådan nogle ting man arbejder på, og det er jo beslutningsstøtte inddirekte, fordi der er nogle præparerater man ikke kan give, hvis man har allergi over for x,y og z, ikke, og hvis det ligesom er registreret at du har det, så vil den jo poppe op og sige at det her er ikke en smart ide. Så kan der være nogle grunde til det så du kan overrule det, men ellers, så det er noget man arbejder med i digitaliseringstrategien, og også konkret og udmønte, så det kan godt være at det var interessant for jer at kigge på, den ligger online, altså den kan man bare hente.

32:50 Mikkel: Kan du komme i tanke om forskellige opfattelser patienter kan have, som kan have en effekt på de her, øhhm, en patients tillid i forhold til de råd eller beslutninger der bliver taget?

33:06 IP: Ja, altså det, øhh, altså, men det er jo ikke noget jeg sådan har erfaring for men altså det er jo bare sådan noget almindelig common sense, altså, Åge nede fra bodegaen ved jo godt selv hvad han fejler og hvad der skal ske, ikke. Så det der med at følge, altså de retningslinjer som man bliver præsenteret for, det afhænger jo rigtig, altså jeg har lavet nogle undersøgelser, med og de ligger også sådan, øhh, inde på DACHIs hjemmeside, sådan nogle technical report,

som viser at befolkningen generelt har rigtig stor tillid til, øhh, til det sundheds-IT, altså, hvis vi undersøger, men altså til IT i sundhedssektoren, jeg tror at over 60

34:04 Mikkel: Hvor sagde du de lå?

34:05 Øhh, inde, DACHI, altså D-A-C-H-I.aau.dk og der under publikation og technical reports, er der så et underfladeblad der hedder og jeg tror også at, hvis i kigger på mig på VBN, så vil i også finde, vi har kun lavet en rapport for 13' og 15', men jeg har også data fra 17'. Øhhm. Og vi skal igang med en ny her fra 19', men generelt har borgerne stor tillid. Altså, måske større en journalisterne, ikke.

34:44 Mikkel: Ja. Og så lidt omkring den her autoritet som du også snakker om, at lægen egentligt er højest i hierarkiet, øhh, tror du også det har en effekt på når de, den måde de har tillid til de forskellige roller, altså...

35:05 IP: Patienterne?

35:05 Mikkel: Ja. Hvor meget den her autoritet spiller ind på deres tillid til de råd der bliver givet, eller de beslutninger der bliver taget?

35:12 IP: Nej, nej det tror jeg ikke man kan sige noget generelt om, fordi det er jo sygeplejersken der ser meget. Så mange vil jo have tillid til dem, fordi det er jo dem de kan se, der laver arbejdet. Altså, men det tror jeg ikke man kan sige noget generelt om, eller jeg kan i hvert fald ikke noget generelt om det.

35:26 Mikkel: Nej.

35:26 IP: Men det er der jo også nogle der har målt, altså man laver jo sådan nogle, øhh, tilfredshedsundersøgelser hvert år, det kan man jo gå ind og finde på sundhedsdatastyrelse, hvad borgerne mener om den behandling de får, øhhm, og så kan jeg huske engang jeg læste en, det er nogle år siden, ikke også, men altså jeg tror overskriften, det var fra Ringkøbing Amt af, 'de er så søde, men kors, hvor har de travlt', ikke. Altså, men, der er jo noget problematisk, rent videnskabeligt er der jo noget problematisk ved de der undersøgelser, hvor man spørger folk 2 måneder efter de er kommet hjem fra sygehuset, hvad de synes om det da de var indlagt, ikke. Øhhm, hvis de var raske, så synes de det var fedt, det er jo ligesom det vi går op i, ikke. Og ikke så meget om de gik med klipklapper eller ej, altså, og kan måske ikke helt huske, hvor elendigt maden smagte, når der er gået to måneder, ikke, men det kan godt være at man er blevet bedre til at lave det, men altså generelt er folk jo meget positive, især hvis de bliver helbredt ikke.

36:25 Mikkel: Ja det er rigtigt.

36:31 Mikkel: Ved du om der gøres noget for ligesom at beskrive beslutningstagningen for patienten, sådan i forhold til behandlingen osv.?

36:43 IP: Nej, eller jo, eller nej. Altså der gøres jo meget for at fortælle patienterne, hvad det er der skal foregå, men jeg tror ikke man skriver, hvem der beslutter det i de pjecer man udleverer til patienterne, siger man jo bare, at så kommer du ind, og så vil der komme en læge og tale med dig, og så vil man gøre sådan og sådan, og så kommer der, og så skal du bedøves og så skal du osv. osv. osv. Men det er ikke mit indtryk at man gør meget ud af at sige hvem der træffer beslutningerne omkring det, det er ligesom bag tæppet, ikke.

37:11 Mikkel: Heller ikke sådan baseret på den her variabel, så gør vi det her, eller, er det mere dit indtryk at de siger, 'så gør vi det her nu'? Hvis nu de skal opereres eller sådan noget.

37:20 IP: Altså nej, der, altså jeg kan jo bare tale ud fra mig selv. Jeg har en kunstig hofte så jeg er blevet opereret en gang. Og man er jo igennem mange steps, altså først så er man ved egen læge og man har ondt og man gør ved og man bliver berørt og og så er der så meget indtil man sådan når derinde, hvor man står i sokker, ikke. Øhhmm, og så vil de jo, altså jeg kan huske da jeg blev opereret i hoften så havde vi, altså vi var tre der skulle opereres næste dag og der havde vi et møde med den læge der skulle operere os. Og han havde simpelthen protesen med og han viste os hvordan den øhh, hvad det var han propede ind i os og hvordan det fungerede, men det var jo ikke sådan så at hvis, at jeg kunne sige at jeg ville hellere have en anden protese. Altså fordi der var jo ikke et udvalg, altså det er jo ikke en skobutik. Øhh, og, men, men selve beslutningen om at operere mig, den er jeg jo blevet præsenteret for og spurgt om jeg så havde lyst til at blive opereret, altså hvor de siger at 'vi kan se at brusken inde i dit gamle led', jeg har engang haft en cykelulykke og sådan noget, ikke, men 'der er ikke mere du går knogle mod knogle, det er derfor du har ondt og det siger *knirk knirk*, hver gang du rejser dig, og det bliver ikke bedre, det vi kan tilbyde det er at du får en kunstig hofte', så er det jo op til mig at vurdere om jeg vil have det eller ej. De siger ikke at vi kan tilbyde dig at du kan få en kunstig hofte, a, b, og c, eller d, og det kan godt være at man gør det i dag, men nu er det jo også tyve år siden at jeg blev opereret, ikke, men altså, fordi der findes flere forskellige typer i dag, altså der er noget der hedder en sportshofte som man giver til folk der skal dyrke sport og sådan noget, hvor kuglehovedet er større, altså, og der kan jo godt være en diskussion af, øhhh, altså jeg har en veninde som har en dårlig hofte og som skal have en ny fordi, altså hvad ved jeg, og hun sagde, 'jeg skal altså have sådan en sportshofte igen, de skal ikke sige at fordi jeg er blevet 67 så skal jeg have sådan en anden en, fordi så kan jeg ikke kravle ned under mit

væv og gøre ved, altså, øhh, så på den måde så vil man jo, men, men, det kræver jo at man selv interesserer sig for det.

39:18 Mikkel: Ja.

39:18 IP: Jeg ved ikke om det er svar, men, øhh, det da et bud.

39:21 Mikkel: Ja. Øhhmm. Tror du patienter opfatter de her råd forskelligt fra hvem der giver dem, og hvem der tager beslutningen, hvis du det for eksempel var en læge eller en sygeplejerske, tror du der er forskel på, hvordan de opfatter de beslutninger og de råd?

39:49 IP: Altså om det er en læge der siger at nu skal du holde op med at ryge eller om det er en sygeplejerske der siger nu skal du holde op med at ryge?

39:55 Mikkel: Ja det er et eksempel ja.

39:57 IP: Det, det, øhh, det vil jeg ikke kunne sige, fordi det afhænger jo af konteksten, altså øhh, og det afhænger jo af, hvem er patienten. Er det Åge nede fra bodegaen, så ville han give fingeren til lige meget hvad de siger, eller og måske ville han lytte mere på lægen, det ved jeg ikke, men, men ellers så kommer det jo an på hvad kontekst. Altså hvis de ligesom siger jeg har taget et røntgenbillede og jeg kan se at du har pletter der og det er begyndende lungekraeft og måske du skulle holde op med at ryge og så kan vi osv., øhhh, der vil det jo typisk være lægen der står og viser røntgenbilledet, men så er det jo ikke fordi at det er en læge, men så er det jo fordi at den viden du får sammen med, altså hvis det var en sygeplejerske der sad og viste dig det, så ville du også lytte på det, på en anden måde end bare at få at vide at det er smart at holde op med at ryge, fordi det er sundhedsskadeligt ikke. Så men jeg tror ikke at det betyder særligt meget om det er den ene eller om det er den anden. Det, øhhm, nej, det ved jeg ikke, hvor i har fra det der. *haha*. Men det må i have læst et eller andet sted.

40:59 Mikkel: Ja, for eksempel, nu siger du så det her med informationen de så giver, det kan være en af de faktorer der spiller ind på, om de, om de, øhhmm, vil følge rådet for eksempel.

41:11 IP: Altså jeg tror at det betyder noget, den måde man kommunikerer med patienter på, altså, om man siger at, øhh, altså, om man giver sig, tid til at forklare at, at de videnskabelige undersøgelser der er lavet, som involverer, øhhh, der er noget der hedder flostrupundersøgelsen, der er sådan nogle store cohorte undersøgelser, ikke, hvor man har fulgt 50.000 eller 20.000 mennesker i 50 år, ikke, og det viser at dem der har røget, *banker i hånden*, sammen de får de og de sygdomme, og dem der ikke har røget de får ikke de og de sygdomme. Dem der holdt op efter så og så mange år, og så mange du har røget, der sker det og det. Det er jo det at sætte sig ned og forklare det, mere end hvem der forklarer det, altså, det tror jeg på, det er det med at fortælle, altså, give argumenterne og, give, give, give, forklare den, altså, på en måde så det er til at forstå.

42:08 Mikkel: Ja.

42:08 IP: Det tror jeg det betyder mere end, end øhh, end hvem det er og hvad baggrund de har, og der kan man sige at der er sygeplejerskerne måske i virkeligheden bedre til at forklare det, fordi de er ligesom mere i øjenhøjde med, med patienterne, ikke, og ved hvad der skal til, hvor lægerne mere ligesom, øhh, skal respekteres af deres kollegaer og være fine på den, ikke. Fordi de, øhhmm, har et andet sprog, de taler med hinanden som ikke er lægmandssprog, ikke.

42:30 Mikkel: Ja.

42:30 IP: De kommer nemt til at tale som om at de sidder overfor en anden læge og så fatter man ikke en brik af hvad det er de siger, ikke, og det er jo ikke ond vilje, men det er fordi det er den måde de, ja.

42:41 Mikkel: Tror du, ved du om der er andre faktorer der også kunne spille ind, ud over den her måde man kommunikerer på?

42:49 IP: Altså i forhold til og, det ved jeg ikke helt hvad det er i forhold til.

42:51 Mikkel: Ja i forhold til at en patient skal følge dit råd, som kom fra personalet.

43:00 IP: Ja, altså fordi det jo, det der også spiller ind det er det patienten kommer med, altså som jeg siger med Åge nede fra bodegaen, altså hvad har, hvilke kompetencer har man selv, altså hvordan er din health literacy, altså dit, øhh, literacy, health literacy, og snakker vi også om eHealth literacy.

43:19 Mikkel: Ja, okay.

43:19 IP: Det er sådan nogle begreber, altså, evnen til at læse og skrive ikke, og lix-tal osv. og health literacy er forståelsen for din egen sygdom og den, øhh, hvor der kommer fra og hvad du kan gøre ved det ikke, og eHealth literacy, det er når vi bruger IT til at mediere, det er alle jeres algoritmer og hvad ved jeg, altså evnen til at forstå, øhh, og det der også betyder noget, og det er jo så det jeg selv forsøker i, det er hvordan designer du den teknologi, som du gerne vil bruge til at forbedre patienternes opfattelse. Altså hvis du sidder, øhh, unge smarte datalogerude på Cassiopeia og designer det, som vi synes at det skal designes, ja så får du også et system der kan bruges af unge dataloger som har fattet hvordan det handler, men hvis du gerne vil have Åge nede på bodegaen og hans kammerater til at ændre livsstil så er du altså nødt til

at designe det sammen med dem. For at forstå hvad det er der betyder noget i deres verden, og hvad der kan få dem til at, altså det er så simpelt, men alligevel meget meget svært. Vil jeg sige. Og der snakker vi om, altså der er en god artikel som øhh, omkring det der med design af systemer, som måske, som nogle kollegaer har skrevet, øhhm, Showel og Turner, og de snakker om PLU og DDD, i forhold til design af IT-systemer. PLU det er people like us, og DDDs det er Disconnected Disempowered og, og, hvad hedder det, dis, ah ja, altså folk der ligesom ikke er med på, øhhm, beatet ikke, men som er dem der har de største sundhedsproblemer, så altså Turner hedder han, hvis du skriver Turner PLU og DDD så tror jeg der sker noget, Poul Turner, han er godt nok andenforfatter, jeg kan ikke, den anden hedder Showel, men jeg kan ikke huske, hvordan man staver til det, så, men men det sætter ligesom fokus på det der med at den teknologi vi designer, øhhh, måske ligesom ikke er designet til den målgruppe vi gerne vil fange og det er jo egentlig det i sidder og snakker om, ikke. I forhold til kommunikation og sådan noget, ja.

45:21 Mikkel: Øhh, hvordan tror du så i forhold til sundhedspersonalet og deres indtryk af beslutningsstøttesystemer, øhh, og hele det her med at have software til at.

45:31 IP: Altså det ville jeg jo spørge de sundhedsprofessionelle om.

45:34 Mikkel: Ja.

45:39 IP: Men altså prøv at læse om det, altså det ville, og så gå over og snak med dem på sygehuset, eller gå over og snak med dem på klinisk institut, der kan i jo fodre svin med sundhedsprofessionelle derovre. Der er en uddannelse der hedder klinisk videnskab og teknologi ovre på HST på sundhedsteknologi, som er en uddannelse for sundhedsprofessionelle som så læser, altså som læser kandidat der er også folkesundhedsvidenskab, hvor der også er massere af sygeplejersker og fyser og alt muligt, der læser en kandidat, så hvis ikke i vil ind på, så kunne i gå derover og snakke med dem om, om de der spørgsmål, ikke, og så kunne i måske også, hvis i vil finde nogle læger, så kunne i gå over og finde nogle specialestudenter, som har væretude, fordi de kommer jo ud allerede, altså de kan jo arbejde som læger inden de er færdige, de kan komme ud i sådan noget halløj, ikke. Og så prøv at spørge dem, de ville være meget bedre til at svare på det.

46:28 Mikkel: Ja.

46:30 IP: Jeg gætter.

46:32 Mikkel: Jamen det var sådan set også de spørgsmål vi havde.

A.3 Analysis

This section contains the analysis of interview I. The analysis is seen in **Table A.2**

Table A.2: Analysis of interview I concerning the healthcare professionals

Interviewsekvens	Meningskondensering	Analyse og fortolkning
00:43 - 01.54	<p>Uddannet teknologi- og samfund-splanlægger. - Fokusområde infrastruktur.</p> <p>-</p> <p>20 års erfaring med sundhedsinformatik - sidste 10 år specialiseret i brugerinddragelse i forbindelse med design af IT.</p> <p>-</p> <p>Arbejdet med metoder for evaluering af sundheds-IT, specielt EPJ.</p> <p>-</p> <p>Har studeret sundhedspersonales arbejdsgange, med fokus på læge-sekretærer.</p>	<p>Personen har ekspertviden inden for sundhedssektoren. Gennem 20 års forskning i sundhedsinformatik, IT i sundhedssektoren.</p> <p>-</p> <p>Personen har gennem sin forskning dannet sig et overblik over sundhedsområdet. Især i forhold til strukturer og arbejdsopgaver i sundhedsområdet. Men også i forbindelse med brugerinddragelse og design.</p> <p>-</p> <p>Har ikke været en del af sundhedspersonalet ud over felt-, interview- og observationsstudier.</p>

Continued on next page

Table A.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
04:28-17:12	<p>Lægerne er de videnskabeligt funderet, der arbejder ud fra evidens (Positivistisk). Sygeplejerskerne er hermeneutiske i deres tilgang og arbejder ud fra det hele menneske, og omsorgen/plejen af mennesket når det er blevet behandlet.</p> <p>-</p> <p>Det er lægen der har magten på sygehuset. Lægen bestemmer da de har ansvaret</p> <p>-</p> <p>Lægesekretæren og lægen har en alliance, da lægesekretæren har været kvalitetstjekker på hvad lægen bestemte.</p> <p>-</p> <p>Lægesekretæren har en faglig forståelse.</p> <p>-</p> <p>Sundhedsjournalen har gjort at lægesekretærerne ikke længere kan hjælpe lægen.</p> <p>-</p> <p>Sygeplejersker arbejder uden sekretærer</p> <p>-</p> <p>En sygehusafdeling har to ledere, en overlæge og en oversygeplejerske</p> <p>-</p> <p>Lægesekretæren styrer meget af det sundhedsadministrative men er ikke anerkendt som en del af afdelingsledelsen</p> <p>-</p> <p>Djøferne prøver at styre hospitalerne fra regeringens side (Værdibaseret ledelse)</p> <p>-</p> <p>Arbejdssopgaver er meget kontekstspecifikke. Medicinske afdelingerne bruger meget tid på at finde ud af hvilken sygdom det er. Kirurgiske bruger mest tid på at behandle sygdommen (Operationer, osv.)</p> <p>-</p> <p>Sygeplejersker er specialiserede i forskellige områder i dag.</p> <p>-</p> <p>Der findes mange typer/roller i sundhedsvæsenet i dag.</p> <p>-</p> <p>Fysioterapeuter og diætister har ikke lige så meget autorisation, og det har været et punkt ift. til adgangen af patientjournaler.</p>	<p>Læger har et positivistisk syn på behandling af patienter. Læger arbejder ud fra videnskabelig evidens og er videnskabeligt funderede.</p> <p>-</p> <p>Sygeplejersker har et holistisk syn på patienter. Sygeplejersker er hermeneutiske og arbejder ud fra det hele menneske. De tager sig af omsorgen og plejen af mennesket når de er blevet behandlet.</p> <p>-</p> <p>Lægesekretærer er kvalitetstjekker på lægens arbejde og styrer meget af det administrative</p> <p>-</p> <p>Der findes overlæger, oversygeplejersker som leder afdelinger. Disse kan muligvis have en anden/ekstra autoritet overfor patienter</p> <p>-</p> <p>Arbejdssopgaver er meget kontekstafhængige, og mere information skal bruges her hvis et medicinsk/kirugisk scenarie skal bruges i experimentelt forsøg.</p> <p>-</p> <p>Der findes mange roller; Laboranter, Fysioterapeuter, Ergoterapeuter, Diætister, Portører. Ikke alle arbejder med patienter selvom de har en rolle på et sygehus. Patientkontakt er vigtig i forhold til råd, da man ellers ikke vil tillægge personerne autoritet/vinden omkring sygedommen.</p> <p>-</p> <p>Diætister og fysio/ergoterapeuter er eksempler på roller med mindre autorisation end fx. læger og sygeplejersker, og som ikke nødvendigvis er ansat på sygehuset.</p>
04:28-17:12 (fortsat)		Continued on next page

Table A.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
18:12-30:01	<p>Læge og sygeplejerske fungerer som et team, ift. til stuegang.</p> <p>-</p> <p>Patienter har mere kontakt med sygeplejersker da de har mere med generel pleje at gøre.</p> <p>-</p>	<p>Selvom læge og sygeplejerske fungerer som et team til stuegang, så er patienter oftest mest i kontakt med sygeplejersker. Her er det den hele pleje der er i fokus og ikke kun behandlingen af sygdommen (Holistisk).</p> <p>-</p>
18:12-30:01 (fortsat)	<p>Organiseringen er kontekst specifik. Bestemte afdelinger har deres måde at gøre ting på.</p> <p>-</p> <p>En læge kan have bestemte patienter han/hun har opereret i, og det samme med sygeplejersker</p> <p>-</p> <p>Lægens råd bliver vægtet højere, pga. at være en specialist.</p> <p>-</p> <p>Mange patienter bruger dog også internettet (Sundhed.dk/Netdoktor.dk) som informations/rådgiver. Relationen patient/læge har ændret sig fra i gamle dage.</p> <p>-</p> <p>Sygehuset opfordrer patienter at sende spørgsmål så de husker at få dem afklaret.</p> <p>-</p> <p>Læger kender også deres begrænsning.</p> <p>-</p> <p>Mange ting er baseret på erfaringer, som vi endnu ikke ved hvorfor fungerer/ikke fungerer.</p> <p>-</p> <p>Egen erfaring: Uddannelsesbaggrund og hvorvidt de er ansat på sygehuset spiller en rolle for hvilket råd der skal lyttes til.</p> <p>-</p> <p>Ville lytte til en diætist mere end læge, hvis diætisten havde uddannelse med baggrund i ernæring.</p> <p>-</p> <p>Lægen har ansvaret og står for beslutningerne</p> <p>-</p> <p>Konferencer afholdes i dagligdagen, hvor de diskuterer hvilke patienter der er kommet ind og hvor mange.</p> <p>-</p>	<p>Lægen har ansvaret, men uddelegerer tit ansvar til sygeplejerske (PN eksempel) Lægen er specialisten, hvilket patienter godt er klar over. Det er dem der opererer/beslutter medicin -> Dermed vil lægens råd blive vægtet højere.</p> <p>-</p> <p>Patienter bruger mere og mere internettet til at undersøge symptomer. Findes der litteratur på dette? Hvor dan vil en person vægte deres egen sygdomsundersøgelse mod det lægen siger?</p> <p>-</p> <p>Vægtning af råd har baggrund i hvilken uddannelse fagfolk har, diætisten kan godt være mere pålidelig i ernæringsrådgivning end en læge.</p> <p>-</p> <p>Lægen (Intensiv) anvender en lille bog/notesblok til at bidrage til beslutningsprocessen/diagnos-ticering. Men patienter bliver også involveret i beslutninger om behandlinger.</p>

Continued on next page

Table A.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
18:12-30:01 (fortsat)	<p>På intensiv afdeling anvender læger forskellige journaler. Dokumentationsjournalen som skrives ind af sekretæren og skal bruges offentligt, og deres private journalliste, hvor ting bliver noteret. Den private bruges til at bidrage til beslutningsprocessen da man på intensiv har brug for at have opdaterede oplysninger konstant.</p> <ul style="list-style-type: none"> - 	
30:36-43:19	<ul style="list-style-type: none"> - Lægen kan ordinere medicin og skrive "PN". Dette betyder efter behov, og giver sygeplejersken ansvaret for at styre det. - Lægen uddelegerer langt hen ad vejen ansvaret til sygeplejersken. - Det handler også om patientens egen erfaring ift. til beslutninger om behandling. - "Det er jo ikke matematik" <p>Digitaliseringsstrategien for sundhedsvæsenet fra 2018 beskriver, hvad der arbejdes på i forhold til digitalisering og beslutningsstøtte.</p> <ul style="list-style-type: none"> - Der arbejdes med at lave bedre beslutningsstøtte i forhold til præparerter, som ikke må tages sammen fordi de enten ophæver hinandens virkning eller gør noget andet, når de tages sammen. - 	<p>Der er nye beslutningsstøttesystemer på vej. Disse kan findes i digitaliseringsstrategien for 2018.</p> <p>Litteratursøgning viser at der er tre nye beslutningsstøttesystemer på vej. Det drejer sig om beslutningsstøtte til ordinering af medicin, krydseffekter og allergier, teknologi til prædiktion, beslutningsstøtte og læring og beslutningstøtte til opsporing af ældre, gennem digital og systematisk registrering af deres sundhedstilstand.</p> <p>Ingen erfaring for beslutningsstøttesystemer, men trækker på sund fornuft. Det kommer meget an på personen. Findes der bestemte demografier der vil være mindre tilbøjelige til at stole på støttesystemer/AI?</p>

Continued on next page

Table A.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
30:36-43:19 (fortsat)	<p>Der arbejdes ydermere på bedre beslutningsstøtte i forbindelse med allergier, eksempelvis overfor hø, penicillin eller noget tredje. - I den forbindelse arbejdes der på et centralt system, CAVE, til disse data. - I dag er der ikke noget automatik om det og der findes heller ikke en central, der arbejdes derfor på at oprette en national database. - I dag spørger man patienter om de selv kender til nogle allergier. - Dette fungerer også som indirekte beslutningsstøtte da der er nogle præparerter man ikke kan få, hvis man har en bestemt type allergi. - Allergi for en bestemt type medicin kunne eksempelvis poppe op på skærmen, hvis det er problematisk for en bestemt type medicin.</p> <p>- Personen har ikke meget erfaring inden for beslutningsstøttesystemer specifikt. - Personen har heller ikke meget erfaring med patienters opfattelse af og tillid til de råd der modtages.</p> <p>- Det er common sense, hvem man stoler på, i forhold til beslutningsstøttesystemer. - 'Åge nede fra bodegaen' ved godt selv hvad han fejler og hvad der skal ske.</p> <p>- Personen har lavet undersøgelser om befolkningens tillid til sundheds-IT. Disse tekniske rapporter kan findes på DACHI.aau.dk. - Der er spurgt over en årrække. - Der er tilgængelige rapporter fra 2013 og 2015 og der er data fra 2017. - Undersøgelsen for 2019 skal til at påbegyndes.</p>	<p>- Rapporter omkring befolkningens holdning til sundheds-IT findes. Befolkning har stor tillid til sundheds-IT, over 60 % mener det vil give bedre behandling inden de næste 3 år. -> Stemmer dette overens med rapporterne?</p> <p>- Intet generelt kan siges om patienter og tillid til autoriteter. Nævner tidligere at lægens råd vil vægtes højere, men mener ikke at autoritet er effekten, men i stedet specialiseringseffekten.</p> <p>- Patienter bliver informeret om processen, men ikke hvem der beslutter ting, hvornår i processen. Dette er dog kun ud fra egen erfaring at det bliver nævnt. Som nævnt tidligere kan dette være meget kontektspecifikt (Type afdeling, region, osv). Eventuelt kan det undersøges hvordan/hvor meget en specifik afdeling involverer en patient i processen gennem observationer, etc.</p> <p>- Sundhedsdatastyrelsen har lavet tilfredshedsundersøgelser, men disse er ikke særligt videnskabelige. Der er derfor ikke særlig høj validitet i disse undersøgelser.</p> <p>- Flere og flere valgmuligheder opstår for patienter og behandling. Dette kan evt. også være en grund til at folk selv undersøger sygdommen på nettet omkring alternativer.</p> <p>- Patient er involveret i processen omkring valg af behandling. Hvis ingen operation ønskes er dette muligt, men ift. til hvem der diagnosticerer/beslutter er patienten ikke involveret.</p>

Continued on next page

Table A.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
30:36-43:19 (fortsat)	<p>Befolkningen har generelt riktig stor tillid til sundheds-IT. Over 60 % af befolkningen mener at det vil give dem en bedre behandling inden for de næste tre år. - Befolkningen har måske større tillid til sundheds-IT end journalisterne.</p> <p>-</p> <p>Man kan ikke sige noget generelt om at patienter stoler mere på nogle autoriteter inden for sundhedssektoren.</p> <p>- Patienten ser sygeplejersken oftest, hvorfor mange har stor tillid til dem, fordi patienten kan se at det er dem der laver arbejdet. - Personen mener ikke at hun kan sige noget generelt om dette.</p> <p>-</p> <p>Der er lavet tilfredshedsundersøgelser af sundhedsdatastyrelsen, men disse er ikke særligt videnskabelige.</p> <p>-</p> <p>Der gøres meget for at fortælle patienterne, hvad der skal foregå når de indlægges/bliver syge. Men der beskrives ikke, hvem der tager, hvilke beslutninger i forløbet. - Patienter får udleveret pjecer, men disse fortæller kun om forløbet, og ikke om beslutninger. - Der gøres ikke meget for at sige, hvem der træffer beslutningerne 'bag tæppet'.</p> <p>-</p> <p>Personen kan kun tale ud fra sig selv i forhold til patienters information om beslutningsprocessen, - Hun er selv blevet opereret på hospitalet.</p> <p>-</p> <p>Når man skal opereres er man igen nemt mange steps - man starter for eksempel ved egen læge når man har ondt og slutter af med at stå inde på sygehuset i sokker.</p> <p>-</p>	<p>-</p> <p>Mener ikke at patienter opfatter råd forskellige fra læge til sygeplejerske. Nævnte tidligere at lægens råd var vægtet mere pga. specialisering. Men nævner nu at kommunikationen er vigtigt, og ikke om det er en læge eller sygeplejerske der kommer med rådet.</p> <p>-</p> <p>Måden hvorpå kommunikationen forekommer er vigtig. Argumenter skal fremføres baseret på fakta, og ved brug af undersøgelser og evidens. Patienten skal forstå hvad der sker, og det betyder mere end hvem der giver informationen.</p> <p>-</p> <p>Sygeplejersker er mere i øjenhøjde med patienter. Lægerne bruger ofte et sværere sprog, dog ikke med vilje. Dette kan resultere i sygeplejesker værende mere forståelige for en patient.</p> <p>-</p> <p>Tre faktorer der kan spille ind på patienter og rådgivning om sygdomme. Literacy, Health Literacy og E-health Literacy</p> <p>-</p> <p>Literacy. Evnen til at læse og skrive.</p> <p>-</p> <p>Health literacy og E-health literacy. Health literacy er forståelsen for egen sygdom (Årsag og behandling), mens e-health literacy handler om forståelsen for sundheds-IT. Hvilken litteratur findes der på E-health literacy? Er der relevant litteratur med overvejelser til experimentelt forsøg.</p> <p>-</p>

Continued on next page

Table A.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
30:36-43:19 (fortsat)	<p>Inden operation er der et møde med lægen. Lægen har protesen med og viser, hvordan den virker, og hvor den skal hen. Men det er ikke muligt at vælge frit, som i for eksempel en skobutik. – Det er muligt at der er flere valgmuligheder i dag. Der findes for eksempel en sportshøfte til folk der dyrker sport – For 20 år siden var det ikke et valg. – For at være en del af besluningsprocessen kræver det at man selv interesserer sig for det.</p> <p>–</p> <p>Der præsenteres for patienten om de vil opereres eller ikke. – Patienten for situationen forklaret. – Så er det op til patienten selv at vælge om de vil det.</p> <p>–</p> <p>Kan ikke siges om patienten opfatter råd forskelligt afhængigt af om det er fra læge eller sygeplejerske. – ‘Åge nede fra bodegaen’ ville ikke lytte til nogle af dem. – Han ville nok lytte lidt mere på lægen. – Det er meget kontekstbestemt.</p> <p>–</p> <p>Der vil ofte argumenteres på baggrund af resultater fra nogle tests. – Det er typisk lægen der præsenterer disse resultater. – Det er ikke nødvendigvis fordi det er lægen der giver rådet, men fordi det er ham der præsenterer resultaterne fra en test. – Hvis det var en sygeplejerske der viste resultaterne ville der også blive lyttet til det. – Det betyder ikke særligt meget om det er den ene eller den anden der giver råd.</p> <p>–</p> <p>Meget af tilliden ligger i at argumentet fremføres på baggrund af fakta. – Det betyder mere end at få at vide at man bare skal ændre sin livsstil, for eksempel.</p>	Det er problematisk at sundhedsteknologier ikke bliver designet og udviklet med de reelle brugere. Dette belyses eksempelvis i et studie fra Showell & Turner, hvori de beskriver, at der ofte designes til “people like us” og at dette skaber “disempowerment, disengagement og disconnectedness”. Især for de mest relevante brugere.

Continued on next page

Table A.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
30:36-43:19 (fortsat)	<p>- Det betyder noget, den måde der kommunikeres med patienten på. - Det betyder noget, hvis man kan pege på store videnskabelige undersøgelser. - Det handler om at sætte sig ned og forklare på en måde så patienten kan forstå det. - Det betyder mere end hvem og hvad baggrund de har, dem der præsenterer rådene.</p> <p>- Sygeplejersken er måske bedre til at forklare, fordi de er i øjenhøjde med patienten. - Sygeplejersken ved, hvad der skal til. - Lægerne skal respekteres af deres kollegaer og være fine på den. - Lægerne har et andet sprog, når de taler med hinanden som ikke er lægmandssprog. - Lægerne taler nemt som om de sidder overfor en anden læge, og så forstår patienten mindre. - Lægerne gør det ikke med vilje.</p> <p>-</p> <p>En anden faktor der spiller ind i forhold til tillid af råd er patientens kompetencer. For eksempel spiller literacy, health literacy og e-health literacy ind.</p> <p>-</p> <p>Literacy er evnen til at læse og skrive. (lix-tal osv.)</p> <p>-</p> <p>Health literacy er forståelsen for ens egen sygdom. Hvor den kommer fra og hvad der kan gøres ved det.</p> <p>-</p> <p>E-health literacy handler om algoritmer og forståelsen af disse. Altså evnen til at forstå teknologi og algoritmer.</p> <p>-</p> <p>Unge dataloger der designer et sundheds-IT-system, vil ofte designe et system, som unge dataloger kan bruge. - Det er vigtigt at designe med den egentlige bruger af systemet, for eksempel 'Åge fra bodegaen'. - Det handler om at forstå, hvad der betyder noget i brugerens verden.</p> <p>-</p> <p>Der er en videnskabelig artikel som snakker om at designe sundhedssystemer til dem der har mest behov for dem. - Den er skrevet af Showell og Turner. - Der snakkes om PLU (People like us) og DDD (Disempowered, disengaged, disconnected). - Der sættes fokus på at den teknologi der designes ikke er designet til målgruppen.</p>	
30:36-43:19 (fortsat)		

Continued on next page

Table A.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
45:21-45:39	<p>Har ikke erfaring med sundhedsprofessionelles holdning til beslutningsstøttesystemer. Foreslår at tage fat i netop de sundhedsprofessionelle.</p> <p>-</p> <p>Oplagte steder at starte kunne være; udannelsen klinisk videnskab og teknologi på HST, uddannelsen folkesundhedsvidenskab, specialestudierende læger eller på sygehuset.</p> <p>-</p> <p>Danske sygehuse er åbne overfor at deltage i universitetsprojekter.</p>	<p>Personen har ikke nok erfaring til at kunne udtale sig om emnet. Det er bedre at benytte andre metoder til at undersøge sundhedspersonalets indtryk af algoritmer, beslutningsstøttesystemer og tiltro til algoritmer.</p> <p>-</p> <p>Der kan laves et eksplorativt studie, hvor man tager ud i felten. Eksempelvis kan der laves felt-, interview-, eller observationsstudie på selve sygehuset. Der kan også laves interviewstudier med lægestuderende, på uddannelsen Medicin, eller sygeplejersker der videreuddanner sig på for eksempel uddannelserne; folkesundhedsvidenskab eller klinisk videnskab og teknologi.</p>

Interview II

B

This appendix contains the interview guide, the transcription, and the analysis of interview II.

B.1 Interview Guide

This section contains the interview guide for the second expert interview conducted during this study. The interview guide is seen in [Table B.1](#)

Table B.1: Interview guide II

Introduktion: Interviewet omhandler din viden om sundhedsområdet og har til formål at kortlægge dområdet forskellige beslutningsstøttesystemer – eksempelvis algoritmer, AI eller statistiske modeller – der er relevante i forbindelse med folks opfattelse af råd og vægtning af disse.

Anonymitet: Dit navn anonymiseres, hvis dette ønskes.

Båndoptagelse: Interviewet optages på bånd og transskriberes.

Tid: Interviewet vil tage maks en time.

Forskningsspørgsmål	Interviewspørgsmål
Hvem er interviewpersonen og hvad er interviewpersonens kvalifikationer til at være ekspert i området? Har interviewpersonen et overblik over sundhedsområdet? Er interviewpersonen en del af sundhedsområdet, eksempelvis ansat i domænet eller forsker deri?	<ul style="list-style-type: none"> • Hvad er din baggrund/uddannelse? • Hvad er dit forskningsområde? • Hvor lang tid har du arbejdet med/i sundhedsområdet? • Hvorfor startede du i sin tid med at forske i dette område? • Har du arbejdet med forskellige personale i sundhedssektoren (Læger, sygeplejesker, fysoiterapeuter, diætister, etc.)? • Har du arbejdet med forskellige algoritmer i sundhedssektoren (Statistiske modeller, beslutningsstøttesystemer, AI, etc.)?
Hvilke rådgivere – fysiske attributter – er relevante i forbindelse med folks opfattelse og vægtning af råd? Hvordan kan disse karakterer beskrives?	<ul style="list-style-type: none"> • Hvilke beslutningsstøttesystemer findes der i sundhedssektoren, i forhold til behandling af patienter? • Kan du beskrive noget der karakteriserer de forskellige systemer/algoritmer? (Hvad går de ud på/Hvad laver de?) <ul style="list-style-type: none"> – Hvad karakteriserer forskellige beslutningsstøttesystemer (for eksempel en algoritme, en statistisk model, AI osv.?) – Benyttes der forskellige beslutningsstøttesystemer af forskellige personer i sundhedssektoren? • Fungerer nogle beslutningsstøttesystemer som rådgivere for patienten? • Fungerer nogle beslutningsstøttesystemer som rådgivere for lægen? • Fungerer nogle beslutningsstøttesystemer som rådgivere for sygeplejersken? • Er det dit indtryk at læger/patienter/sygeplejersker opfatter og vægter råd fra forskellige rådgivere (algoritmer, andre læger, etc.) forskelligt? <ul style="list-style-type: none"> – Kan du beskrive en situation, hvor dette er tilfældet? • Hvem står for beslutningstagningen? <ul style="list-style-type: none"> – Benyttes der teknologi – <i>algoritmer</i> – til beslutningsprocessen? – Står algoritmer for dele af beslutningstagningen? – Kan du nævne nogle situationer, der beskriver eksempler på beslutningstagningen? – Hvem står som ansvarlig for patienten? (Eg. læge, algoritme, sygeplejerske, etc.) – Er patienten en aktiv del af beslutningsprocessen? • Kan du komme i tanke om, forskellige opfattelser personer kan have, som påvirker personens tillid til råd og beslutninger? <ul style="list-style-type: none"> – Er der bestemte opfattelser der er relevante for læger specifikt? – Er der bestemte opfattelser der er relevante for patienter specifikt? – Er der bestemte opfattelser der er relevante for andet sundhedspersonale specifikt? • Gøres der noget for at beskrive beslutningsprocessen til modtageren? <ul style="list-style-type: none"> – Beskrives forskellige algoritmers beslutningstagningen forskelligt? • Hvordan opfatter modtageren forskellige algoritmers beslutninger/råd? <ul style="list-style-type: none"> – Kan du komme i tanke om nogle faktorer, der påvirker modtagerens følelse af tillid til beslutninger og råd?
Hvordan beskrives rådgiveres beslutningstagten for patienter/læger/etc.? Beskrives forskellige rådgiveres beslutningstagten forskelligt? Hvordan opfatter modtageren forskellige rådgiveres processer/beslutninger? Hvad er vigtige faktorer i forhold til at føle tillid til beslutninger/råd?	<ul style="list-style-type: none"> • Gøres der noget for at beskrive beslutningsprocessen til modtageren? <ul style="list-style-type: none"> – Beskrives forskellige algoritmers beslutningstagningen forskelligt? • Hvordan opfatter modtageren forskellige algoritmers beslutninger/råd? <ul style="list-style-type: none"> – Kan du komme i tanke om nogle faktorer, der påvirker modtagerens følelse af tillid til beslutninger og råd?
Hvad er sundhedspersonales indtryk af algoritmer og software? Ses software som hjælpemidler til beslutningstagten? Er der tiltro til at software kan give retvisende råd?	<ul style="list-style-type: none"> • Hvad er sundhedspersonales indtryk af algoritmer og software? <ul style="list-style-type: none"> – Kan du nævne en situation, hvor sundhedspersonalets holdninger kommer til udtryk? – Ses software som hjælpemidler til beslutningstagten? • Er der tiltro til at software kan give retvisende råd? <ul style="list-style-type: none"> – Kan du beskrive, hvilke faktorer, der spiller ind i forhold til om sundhedspersonale har tiltro til software?

B.2 Transcription

This section contains the transcription of interview II. The transcription was made using the oTranscribe website ([Bentley, 2019](#)).

Transcription

00:06 Mikkel: Først vil jeg starte med at høre hvad er din baggrund, og dit nuværende forskningsområde også? 00:13 IP: Ja, såå, jeg har været her siden 97', kom her som post doc, i den her gruppe som den brede gruppe hed "Model baseret medicinsk beslutningsstøtte". Og så har jeg været både post doc og så professor nu. Som professor nu har jeg min undergruppe, som hedder "Respiratory and critical care group". Og det har stadigvæk beslutningstøttesystemer som fokus, men vi rettes ind efter klinisk problemer relateret til lunge funktion, og specielt relateret til intensiv medicin, så de sygste patienter som er intuberet, så har respirationsslange ned i halsen og mekaniskventileret. 01:05 Mikkel: Ja.

01:05 IP: Patienter som er kardiovaskulært støttede og ofte støttede for deres cirkulatorisk aktivitet, så de aller aller dårligste patienter på sygehuset.

01:17 Mikkel: Ja

01:17 IP: Og så, i forskningsgruppe er en af vores største aktiviteter, er vi interesseret for bruge den tunge viden der er omkring matematisk beskrivelse af fysiologi, jeg kender ikke jeres baggrund så godt, men i ingeniør disciplinen er man har, at man kan modellere et system ikke. Og man kan modellere et system med kemi. så, f.eks. vi har kemiske modeller af hvordan blodet fungerer ved dens transport af ilt og kuldioxid og så kan man modellere et system i form af mekanik af deres eftergivethed, deres elastens og den modstand der er osv. Man kan også have modeller af konvektion og af diffusion, der har vi modeller af gasudveksling over lungerne, så lungerne er både mekanisk system og udveksling og en veksling system ikke. Og hvis man har sådan nogle modeller, så kan man både bruge dem diagnostisk ikk, fordi hvis de bliver tunet til patientens data så er parameterne i den model, ting man variere, de bliver en karakteristik af patienten. Så f.eks. hvis man har en lunge der er delvist kollapseret så har blodet gået forbi lungen uden at være involveret i gasudveksling, det kalder man shunned, så hvis man har 40 % shunned så har man en bestemt sygdomstype, som man skal gøre noget ved, så bliver man bedre og så kan man overvåge dig. Men også hvis man har nogle modeller så kan man tune dem ved sendesiden, så kan man lave simulationer ved sendesiden. Så kunne man sige "Hvad vil der ske hvis jeg nu skruede op for tryk, ned for tryk, op for ilt, ned for ilt." Og så kan man simulere hvad responsen på patienten vil være, og hvis man kan simulere hvad responsen vil være så kan man koble beslutningstagning på toppen af den. Og jeg ved ikke om i kender til beslutningsteori, men det er ligesom med spil ikke. Hvis jeg vinder 10 kroner, hvor godt er det?

03:18 Mikkel: Ja

03:18 IP: Jeg kan sætte en straf eller en gain til den ikke. Så hvis jeg skruer ilt op og får mere ilt i blodet, det giver mig en gain ikke. Men til gengæld ilt er giftigt for lungerne, så jeg hvis jeg skruer helt op så får jeg en loss i lungerne ikke. Så kan man afveje hvor meget gain man har i forhold til hvor meget loss man har, og hvis man gør den kvantitativ så bliver det hele et optimiseringssproblem. Så kan man stille en computer til at finde mig de sæt knapper som maksimerer den potentielle, og man har flere domæner den går i ikke. Hvis man giver for meget tryk så kan man skade lungerne af tryk og hvis ikke man giver nok tryk så kan man ikke vaske ud kuldioxid, og så bliver blodet sur, ikke. Så det er den balancering der er ikke, de der balaceringspunkter de varierer meget efter patientens tilstand, hvis lungerne er meget stive. Det forhold mellem at leve tryk og vaske ud CO₂, er helt noget andet ikke, fordi man skal give lav tryk fordi man skal have volume så skal man have højt tryk, men tryk er skadeligt ikke. Så man kan ikke vaske ud så meget CO₂. For hvis lungen er normal, så er det nemt at vaske ud, så kan man lige så godt vaske det hele ud. Det her kompromis, er afhængig af sygdomsforholdet, og her bliver hele det her optimiseringssproblem. Og vi har et system som, som forslag til lægerne ved sengesiden, hvad skal man gøre.

04:41 Mikkel: Er det, det Beacon?

04:42 IP: Det er Beacon ja. Så den kommer fra vores forskning her ikke. Forslag til lægen, hvad man skal gøre, lægen gør hvad de vil. Så måler vi responsen og når patienten responderer så lærer vi fra responsen. Så hvis man, hvis man *Skubber til bord*, og den giver sig, så ved man hvad responsen er ikke, og hvis den holder så ved man den respons, så lidt det samme ikke. Man laver en ping af systemet, hver ændring i intubationen på den non linære system ikke.

05:16 Mikkel: Øhm, i forbindelse med det her projekt, har du så arbejdet med forskellige sundhedspersonale også som en del af det?

05:25 IP: Jo, altså projektet er nu at, øhm, først man har udviklet, firmaet har rettigheder

til ideen, så firmaet har været en spin off, jeg har været medstifter af firmaet, men den lever sit eget liv og giver faktisk penge tilbage til universitetet til vores forskning. Med Firmaet, har udviklet den *Viser tablet* her tablet som ligger på toppen af respiratoren ikke. Så går man ud og afprøver systemet ikke, så kommer man i sygehøst og man designer forskningsprotocol for test, og såå, vores sidste publikation som lige er kommet ud i sidste uge, der er vores første række studier, hvor man har systemet på patient i 6-8 timer, så følger man råd, og så ser man om patienten har bedre tilstand efter rådene. De her giver gode resultater, vi når at, hvad er det man siger øhm, tryk i lungen, vi når at forbedre pH, og formindre ilt problemer osv osv. ikke. Til at køre sådan et studie så skal man arbejde med de her forskningslæger der er i afdelingen, det er den interaktion man har der. Nu har vi været ovre den fase, og den giver lov til CI mærkere systemet, lokalt i Europa ikke, og giver den kredittering og giver den et mærke der siger dette er sikkert. Nu er vi i den fase hvor man laver de store randomiseret kontrol trials og det sker i 8-7 sygehuse i Europa nu, hvor man rekrutterer patienter, og vi har en total patient gruppe, vi skal have 700 patienter, og patienten bliver inviteret og så bliver de randomiseret til at bruge råd fra systemet eller almindelig praksis, og så ser vi hvad er bedst, og så er der nogle kriterier for det. Men det er en helt anden interaktion med personale ikke, fordi nu går vi fra samarbejde med den enkelte læge på afdelingen, at nu skal alle sygeplejesker på den afdeling trænes, for de skal vide hvis, fordi vi har målere inde i slangen der ikke. Hvis den bliver fyldt med væske, så skal de skifte det, så de skal være vant til at kigge på det.

07:38 Mikkel: Ja.

07:38 IP: Og på en intensiv afdeling, der skal bemandes 24/7 ikke, så har man en pænt stor afdeling, så har man 200 sygeplejesker der skal trænes. Det giver en helt anden interaktion og det er firmaet, firmaet betaler sygehøstet for at ansætte forskningssygeplejesker og så har vi haft superusers komme her og vi har trænet 8 superusers til en afdeling, så får de en, en *Viser skilt*, sådan en som de går rundt med i afdelingen og så træner de de andre, og holder øje med at systemet fungerer og at, har ansvar for det daglige løb. Men det giver virkelig en helt anden problemstilling, for nu skal den have den daglige interaktion, nu er det ikke bare mere en person som er venlig mod systemet, som skal snakke med det.

08:34 Mikkel: Og det er også lidt øh, det vi egentlig har tænkt at kigge lidt på i hvertfald, sådan generelt det her med tillid til beslutningsstøttesystemer, øhm, kender du andre beslutningsstøttesystemer som også allerede er i sundhedssektoren?

08:51 IP: Jo, der er nogle af mine kollegaer her ikke. Steen Andreasen har TREAT systemet, som giver antibiotika råd og også har et firma her i byen her. Og Ulrike Pielmeier som derover, hun har et system (Glucosafe) i intensiv afdeling, der rådgiver om ernæring og glukose regulering.

09:16 Mikkel: Ja.

09:16 IP: Men i det store, i den store stil så er der ikke meget, øh, i min egen branche er der to respirator som man kunne kalde beslutningsstøttesystemer, men de er closed loop, så de implementerer bare rådet på respiratoren, det reducerer væsenligt den interaktion der skal være med sygeplejeskerne, fordi den bare gør hvad den skal gøre. Eller læge for den sags skyld, interaktion. Men det er en stor debat der foregår nu, øh, i den domænet som helhed, hvor transparent rådet skal være, hvor forståelig skal rådet være ikke. Men her, den mest promenente bevægelse i AI er machine learning, og groft sagt, man kan måske høre at jeg ikke er den store eksponent af den mode, men man samler mange data, og så laver man en black box model, neurale netværk ikke. Som er ikke mere end en non linæar matematisk korrelation imellem mange variable, og derfor man simuleret noget og så laver man beslutningen udfra det. På den samme måde som jeg lige har sagt ikke. Men man kan ikke sige hvorfor. Og det skulle være udemærket hvis du havde et kontrolsystem, ligesom PXX kontroller og et eller andet ikke. 10:49 Hvor man behøves ikke at sige hvad der er inde i det, men hvad så når det ikke fungerer? Når man skal åbne boksen og forstå det. Det vi prøver, vi prøver at lade folk grave ned til at forstå den, hvordan den fungerer indeni, prøver at gøre den transparent og det betyder at man kan have, man kan have en strategi hvor man siger, skal den her hjælpe lægen ved sengesiden? Skal der være en uddannelsesprocess som støtter og gør folk klogere ved sengesiden, eller skal det støtte en fjernelse af ekspertise og bare gøre ting selv? Det er lidt der, at der er forskellige strategier ikke. I intensiv medicin, er der publikationer der siger at, øh gode undersøgelser at med tiden vi mister anæstesiologer, vi mister ekspertise. Det gør vi med tid ikke. Der bliver et hul, og til gengæld har vi patienter der bliver mere og mere komplekse, fordi vi har mere teknologi til at redde dem, og de fejler flere ting samtidigt, como bedtitis? 11:52 er den store ikke. Så vi både reduktion af ekspertise og så kan man sige, så skal vi have et system der bare kan klare det, eller vi kan sige at vi skal have et system der kan hjælpe med det ved sengesiden, og uddanne og løfte dem der skal nu overtage den ekspertise, det er det vi prøver at gøre. 12:13 Mikkel: Ja. Er det, øhm, dit indtryk at læger eller personalet vægter råd forskelligt som eh, nu idag, fra en algoritme eller fra andre personale?

12:28 IP: Nogen kan ikke lide det og nogen kan lide det agtigt?

12:30 Mikkel: Neeeeej, men sådan, altså, altså, har de mere tillid til det råd der kommer fra deres egen ekspertise eller fra andre end de har fra et beslutningsstøttesystem? Hvordan ser du det idag hos personalet?

12:46 IP: Jeg er ikke sikker på jeg forstår spørgsmålet, øhm, folk har meget forskellige meninger ikke, og meget forskellige fordomme og øhm, det er virkelig op til os at bevise at det fungerer, det er virkelig op, til at få den store støtte igennem som testen reducerer den her ventetid på sygehuset, ventetid på respirator, reducerer dødelighed. Vi skal komme forbi, sådan, jeg synes folk skal være fuldstændig ligeglade med vores teknologi ikke. Der er ikke nogen grund til at hvorfor læger skulle interesser sig for kunstig intelligens, hvorfor skulle de det? De interesserer sig for patienter og det er her vi skal møde dem, med om vi har noget der kan hjælpe deres patienter, hjælpe deres process. Så skal vi snakke om patienter og process, og hvis den her maskine kan hjælpe her så skulle de interesserer, hvis den ikke kan så skal de ikke interesserer sig. Det er ikke i sig selv noget, man køber ikke en telefon for telefonens skyld. Så vi skal snakke med dem på deres egen præmis, synes jeg, og prøve at snakke med dem om konkret patient situation. Hvad skal en XX 14:05 patient når man bestemme hvilken XX måling man skal have? Hvad gør systemet i den situation, hvorfor kan viøre den transparent? Vi kan sige hvorfor, at systemet vil have de modifierer, så de kan opbygge en tro. Vi har f.eks. en funktionalitet hvor, jeg ved ikke om jeg svare på spørgsmålet men jeg snakker lidt og så kan du fiske lidt. Vi har en funktion i vores system hvor man kan sætte grænsen til hvad man vil have råd indenfor, så f.eks. hvis man vil bestemme hvor meget XX 14:41 og man kommer rundt som læge to gange om dagen, og man kan se at den her patient her, at vi skal prøve at holde volumen småt, men ikke alt for småt. Fordi så vil det måske stresse dem lidt for meget. Så kan man sige, "Jeg vil have at vi kan få råd indenfor 400 ml/vejrtræk op til 500/vejrtræk". Så vil systemet give råd ikke. Prøv at presse volumen ned ikke, for det er skadeligt ikke, når den kan. Men når den rammer 400, så siger den "jeg har et råd men det er udenfor grænsen".

15:13 IP: Og så kommer den med et råd der udenfor grænsen, og så kan man have en tynd politik i afdelingen, at når det sker så skal de kalde på lægen. og så har man integreret og støttet den almindelig flow ikke. Man er ikke overtaget, man skal være fleksible indenfor det almindelig flow. Så lægerne kan give ansvar til sygeplejesker, i en hvis grænse, med støtte fra et system og når den er uden for den grænse så skal de kalde igen. Den slags ikke.

15:45 Mikkel: Er der nogen situationer hvor at systemet vil stå for at tage beslutningen selv?

15:53 IP: Nej det gør den aldrig. Ikke vores, men de der closed-loop gør.

15:58 Mikkel: Ja okay.

16:01 IP: Hvis i vil have navnene, så er der en der hedder SmartCare, som er lavet af et firma der hedder Draeger, et tysk firma. Og en anden hedder INTELLiVENT – ASV. Men de er begge to closed-loop, støttesystemer, så ikke rigtigt støttesystemer men reguleringssystemer som sådan, men de er baseret på samme princip ikke, regler eller modeller.

16:26 Morten: Har i kigget på om lægerne ligesom modtager de forskellige systemer, forskelligt, eller sådan, har forskellige holdninger til om det er closed-loop eller om det er støttesystemer?

16:35 IP: Vi har ikke kigget systematisk igennem det har vi ikke, vi har været meget fokuseret på vores egen. 16:43 Morten: Ja.

16:46 Mikkel: Har der været snak om det her med at øhh, hvem der er ansvarlig hvis nu de tager i mod rådene fra beslutningsstøttesystemer, er der noget hvor systemet systemet kan være en del af ansvaret?

17:01 IP: Det er et stor spørgsmål, specielt i USA. Vores attitude er at vi giver råd men vi bestemmer aldrig. Så de ignorere hvad de vil, og det er også de der læringsmekanismer om patient, at man også lærer om patienten hvis de gør noget andet. Fordi det er stadigvæk et pres på systemet.

Hvad den endelige, øhm vi har ikke FDA godkendelse, så firmaet kan ikke markedsføre i USA så vi gør ikke noget i USA. Hvad det endelige vil være i USA, hvis FDA godkendelse kom, det ved jeg ikke. Regulatorisk mæssigt, øhm men jeg kan ikke se det store problem i Europa, øhm regulatorisk, da vi ikke gør noget. Og der er, de to systemer jeg talte om, er i hvertfald SmartCare FDA godkendt. Så det betyder at der er et closed-loop reguleringssystem som er approved ikke, som betyder at der ikke er nogen interaktion men at den bare bestemmer hvad der sker med patienten handlingsmæssigt. Så der er precedence for at det kan lade sig gøre, øhm den kultur i USA er meget "cost" fokuseret, de vil gerne have at patienten er hurtig. Der er også selvfølgelig i sådan nogle systemer en overvågningsaspekt ikke. Nu kan vi se hvad rådet var og vi kan se hvad målingen var. Så man kan gengenerere alle situationer. De patienter vi studerer nu, dem har jeg. Jeg kan gå igennem 14 dage for en patient, for hvilke råd der var, hvorfor kom de råd, hvad var målingen på det tidspunkt, hvad har lægen gjort, hvad har sygeplejeskerne gjort, hvad har de ikke gjort. Så der er en potential dokumentations system som kan bruges

eller misbruges, som nu alt teknologi kan. Så det er en potentiel overvågningskapacitet. Men god voervågning hjælper også på at forbedre kvalitet, dårlig overvågning hjælper med at slå folk på hovedet på problemet. Så teknologien er neutral i den forstand, i mine øjne, men det er klart at anvendelsen ikke er neutral, men er meget politisk præget og kan være både godt og skidt på den front.

19:42 Mikkel: Har i øhm, i forhold til for eksempel patienterne så, har de været med indover eller er de blevet forklaret i forhold til det her med at der bliver brugt et beslutningsstøttesystem?

19:54 IP: I vores forskningsforsøg, der må vi ikke bare gøre hvad vi vil.

19:59 Mikkel: Nej.

19:59 IP: Så i forskningsforsøg ska man have samtykke, men det kan de her patienter ikke give, så det er familien eller lægen med ansvar for protokollen. Med det sagt, så er systemet nu CE godkendt i Europa så den almindelige læge kan bruge det til sit formål i deres almindelig praksis. Og der har været nogen som bare bruger det, så patienten ved ikke noget om at de bruger det. Men det er standard praksis.

Men der er klart regler omkring det her, det er ikke noget vi gætter os frem til.

20:46 Mikkel: Nej nej. Så igen lidt i forhold til det her med at sundhedspersonalet her, du snakkede lidt om at der er forskellige holdninger til deres, øhm, er der generelt tiltro til at den her software der står bag det, kan give retvisende råd?

21:02 IP: Men øhm, generelt tiltro til noget som helst er svært. Folk har så mange meninger.

21:20 Mikkel: Har du et indtryk af hvad der gør at nogen har en negativ holdning og nogen har en positiv holdning, altså hvilke effekter eller hvilke ting der gør? 21:34 IP: Puha, det er en svær en snes jeg, for positiv/negativ holdning kan være på koncept ikke. At nogle folk kna lide teknologi og nogen kan ikke lide teknologi. De kan være på funktion, at nogen synes at rådet er godt og nogen synes rådet er dårligt, i en bestemt situation og i en ikke bestemt situation. Og de kan også være i sådan, småt irritation i anvendelse. I systemet, øhm, nogle modeller passer ikke patienten mere og så beder vi om mere information, som blodgas ikke.

22:11 Mikkel: Ja.

22:11 IP: Og i hvert fald i tidlige versioner af systemet, der bad vi måske om for mange af dem. Det skaber irritation ved sengesiden, og der skal ikke ske for meget irritation før folk giver op. Så det er meget forskellige lag, øhm, at man kan føle sig godt eller skidt. Det kan være helt oppe fra princip af, at man "gider ikke den der ting, jeg ved bedst". Eller "Jeg kan godt lide det, det ser moderne og spændende ud", og begge to er lidt irrelevant på en måde, sådan, lige så meget som at de ikke er velfunderet som meninger. "Jeg kan lide teknologi, jeg kan ikke lide teknologi", det er ikke velfunderet meninger, det er sådan, man skal prøve at trække den ned og snakke om patienter ikke. Hvad var det råd? Hvad sker der hos den patient? Hvorfor synes i det var godt? Hvorfor, man plejer at snakke om det man ikke synes er godt, fordi alt det vi synes er godt snakker vi ikke om. Så der er kraftig bias i hvad man snakker om, fordi ellers kommer der aldrig forbedringer hvis man kun snakker om hvad der er godt. Så man plejer at opbygge historier om, øhm, og det bliver frygtelig kompleks når man er inde i almindelig hverdagspraksis. Langt, langt siden har vi haft en situation(Problem) fordi systemet giver ikke råd hvis patienten er ustabil. Ofte suger man patienten, man sætter et rør ned og suger alt spyt op osv. Og her giver den ikke råd i den periode, Så først når den dedikterer stabilitet så giver den råd igen. Jeg kan huske en situation hvor patienten skal flyttes fra sengen til stolen, og det gør de med, øhm, at trække dem op i sådan et hejseværk, og så bevæge dem over og sætte dem ned igen. Og så bliver de stabile ikke, og så kommer der et råd. Der skal ikke komme et råd midt i sådan en manøvre der, det skal den ikke vel, selvom rådet kunne være fint, så er det upraktisk og forstyrrende. Bare en småt upraktisk og forstyrrende ting, som kan gøre mentaliteten mod systemet forkert. "Ej hvorfor skulle det dumme system komme med et råd nu" agtigt noget ikke. Så skal man have en pause knap ikke, hvor man stopper en periode mens man laver manøvren og ting og sager. Det er meget meget lav praktisk men det er ofte det hvor det fungerer de her irritationer. Ned integrationen med almindeligt flow, og ikke op i den der med "Er råd godt eller skidt?".

Det er meget, meget arbejde et firma skal lave, for det er ikke forskning mere, det er bare lavpraktisk håndtering der skal laves for at være sikker på at den kan integreres i den daglige praksis. Nu kan jeg ikke huske hvad spørgsmålet var, så jeg ved ikke hvis du vil.

25:17 Mikkel: Nåå, ja men det var bare om, den her om, om altså hvilke negative og positive effekter som du kunne.. 25:30 IP: Ja, som jeg siger så kommer den der irritationelle stemning gennem teknologi, så kommer der diskussionspunkter om individuelle patienter eller patientgrupper med folk der kender fysiologi og så kommer der de her diskussioner om småt daglige irritationsemomenter ved daglig praksis. Og systemet er også lidt designet efter de her mennesker ikke. For det vi gør, er at vi bare præsenterer et råd, "Du er her, gør det". Uden forklaring, og hvis man vil have forklaring så skal man grave, og man skal grave efter forklaring

til det niveau hvor man kan forstå det. Så vi prøver også at presse de her folk som lærer deres ekspertise på en eller anden måde ikke. Det er bevidst at vi gør det, også for ikke at irriterere.

26:27 Mikkel: Så er der den her med at, øh, Black box som man ikke vidste helt hvorfor den gjorde som den gjorde osv. Du snakker lidt om at i prøver at lære dem mere, er det så hvordan hele systemet er opbygget og hvad der sker inde med også softwaren osv.

26:45 IP: Vi kan godt køre det, hvis i vil se det?

26:45 Mikkel: Øhh, Ja.

26:50 IP: *Finder tablet frem med systemet på*

28:19 IP: Så i kan forestille jer, hvis i kigger på hjemmesiden så kan i også se det, hvordan respiratorkærmens ligger under og den her *Tablet* ligger på toppen. Gasmålingerne komme så her ikke.

28:32 IP: *Viser hvordan interfacen fungerer og hvordan informationer skal skrives ind.*

28:42 IP: Og såå, så, det er blodstrømningen rundt om i kroppen, i liter pr. minut, det kan man måle eller gætte sig frem til via vægt og højde osv. Lige nu er det bare til modeltuning.

Her er området hvor de vil lade rådet gå indenfor. Så kommer skærmens her med vores målinger, så den her ekspotorisk ilt, ekspotorisk CO₂, blodtryksmeter som viser mig hvor meget ilt i blodet i procent, og det her er CO₂ produktion i vævet, hvis man ved hvor meget CO₂ går ind, og man kan måle hvor meget kommer ud, og matematisk integrerer over tid så kan man måle produktion. Hvis man gør sådan her så har man ilt optagelsen, og så kan man udregne hvor mange kalorier man brænder pga det.

- Flytning af optager-

29:52 IP: Det her respirationskvotienten og er relationen mellem de to. Det er bare noget vævet skal være indenfor 0.7-1, så det giver et kvalitetstjek. Og så kan vi se respirationsfrekvensen er og hvor meget volume de trykker per vejtrækning, hvor meget volume de trækker pr. minut, og hvis vi fjerner de øverste luftveje så er der ikke nogen gasveksling, så øhh, hvor meget luft de bruger per minut, i gasudveksling. Det her hvide mærke er sidste gang vi ændrede på noget, på vores respirator, og den siger vi venter på stabilitet. Den ser meget stabil ud men vi skal have to minutters vindue med stabilitet før vi giver råd. Når to minutter er gået så vil den blive ved, og så vil det begynde at regne.

30:39 Mikkel: Ja okay.

30:39 IP: Og det første den vil gøre så er at have en blodgas, fordi den kan ikke regne alt om kroppen uden at vide noget om blodets status. Så nu beder den om en blodgas. Så taster man værdier ind og så fortsætter den med den blodgas. Godt, så nu har den blodgas, og nu kan den tune den ilt, den har målt her, som fortæller noget om lungerne. Det er tryk og måling i lungerne, eller i ekspiratorisk, den kan måle mekanik egenskaber i lungerne, så du ved blodets karakteristik fordi vi har en blodprøve. Og så nu har den tunet de der modeller så laver den simulering om hvad vi kan gøre, og så laver den en optimisering. Den bedste simulering som viser hvad der er bedst for patienten. Hvis den optimisering er langt langt væk, så tager den først et skridt i den retning, så vi ikke laver for aggressive skridt. Såå, om lidt når den er færdig med sin søgning så kommer den med et råd. Og det kan sagtens tage 2 minutter, det har ingen konsekvens i normal klinisk praksis, da patienten er der. Så vi har massere af computationel tid vi kan arbejde med.

32:04 Mikkel: Men det var mere så øhm, i prøver jo så at lære det her personale hvad der foregår nede bagved.

32:11 IP: Ja. Men det kommer når rådet kommer. 32:11 Mikkel: Okay. 32:12 IP: Først, hvis vi swiper her, så er der ikke nogen simulering klar endnu, når den kommer så kan vi grave lidt i det.

32:30 Morten: Du snakkede om på et tidspunkt, øhm, det med at i lavede de her undersøgelser. Føler du det ligesom er et vigtigt argumenter for at få læger til at bruge det? De her videnskabelige evidens og sådan noget?

32:42 IP: Ja, ikke bare lægerne men det er et vigtigt argument for at administrationen er ville til at købe den. Hvis de sparer en sengedag så er det 25.000 kroner i Danmark. Så, og det koster 10.000 at få sådan et system, så er det god økonomi i det ikke. Jeg ved ikke om det koster 10.000 for at få systemet, det er bare et eksempel. Så det er ikke bare patientens skyld, der er hvert år har en værdi i princip, af leveår, for godt eller ondt ikke. Man vil godt booste sine penge økonomisk, så den evidens er meget vigtig.

33:27 IP: Har vi evidens, så kan man sige at der burde blive brugt det system, så skal man have folk til at acceptere at det er brugt og så det have at folk gør det praktisk.

33:46 IP: Der kommer et råd snart.

33:53 IP: Sådan, så er der et råd okay. Så du kan se der er ikke meget ekspertise i det, den siger ilt er nogenlunde rigtig, XX er nogenlunde rigtigt34:03 , øhm, respirationsfrekvens er nogenlunde rigtig. Og så kan du modificere de der hedder PIP, XX pressure 34:10 . Man kan ventilere et lunge ligesom den her *Klapper*, eller som den her *Viser med hænderne*, hvis

man gör den anden her så er andet ekspertorisk tryk højere og den holder lungerne lidt åben, så hvis man er bekymret for at den kollapsere, så holder man den høj, nogle gange kan man skrue den op så man kan få ekstra volumen i lungerne og bedre gasudveksling. Okay så der er rådet, hvis man gerne vil forstå det råd, så kommer man herovre, og herovre er de der mål man har i klinisk praksis som er mod hinanden. Vi vil gerne give nok ilt, men ikke for meget ikke, så hvis jeg skruer den her ilt op, så bliver der for meget. I kan se den går op der.

34:58 Mikkel: Ja

34:58 IP: Hvis jeg skruer den her ned for patienten, så går den ned og så bliver der for ilt ikke. Så jeg vil gerne have en god balance. Det ligner her omkring ikke, 60 %. Men det er først efter vi har åbnet lungerne lidt ikke, hvis lungerne ikke var åben nok, så kommer vi hened og vi skal hellere op i det område vi var før. Så hvis vi ikke kan åbne lungerne lidt, så kan vi reducere den. Og det er sådan klinisk praksis fungerer ikke, hvis man får en bedre lunge så behøves man ikke have så høj, og ikke så toxic og så er det bedre hele vejen rundt. Så det. 35:49 Morten: Så hvor skal den ligge hen? Skal den ligge ved det grønne eller det gule?

35:55 IP: Man kan ikke altid få den i den grønne, for patienten er ikke rask nok

35:57 Morten: Nej okay. 35:57 IP: Så kan man ikke få lavet ilt, så det skal være mod center så godt som muligt. Og det skal ikke være nogen spikes. Det siger vi til folk, der skal ikke være nogen spikes. Der skal ikke være en retning hvor det peger ud ikke. Så det er den retning der, så lige nu ville den gøre det der. Volumen, jamen volumen giver høj tryk, det giver traume så hvis du drejer på volume så giver du traume ikke. Hvis du skruer ned for volume så kan man ikke vaske CO₂ så bliver blodet sur, så det skal man balancere der. Okay?.

36:40 Mikkel: Ja.

36:40 IP: Nu den her patient, trækker ikke vejret selv. De er dybt sederet. Når de begynder at blive rask så overtager de lidt funktionalitet. Så trækker de lidt selv, og så mærker den at de trækker lidt selv og leverer nogle tryk. Det tryk den leverer, hvis man giver for meget så holder de op med at trække selv, hvis man giver for lidt så kæmper de for hårdt, hvis den holder tryk for dem, så vil musklen, diaframen, den begynder at forsvinde så man får muskelatrofi, og så hvis man trækker for meget så får de stress. Det er den akse der, men den er ikke involveret her fordi patienten ikke trækker vejret. Så i principippet er der lige nu maksimal atrofi.

37:25 Mikkel: Så det ville være en koma patient det her eller hvordan?

37:27 IP: Ja man plejer ikke at kalde dem det i fagligheden, men det er en dybt sederet patient ville man kalde det, som ikke har spontan respiratorisk aktivitet.

37:35 Mikkel: Okay ja.

37:37 IP: Så det er et niveau det her ikke. Sygeplejeske plejer at forstå det, de forstår det her ting. Men lægen kan komme og sige "Det er fint du siger til mig at hvis jeg skruer den her ilt ned, at det bliver værre. Men hvad betyder det?". Nu kan jeg klikke her, og så betyder det at jeg går fra en simuleret artiel XX 38:03, som er 5,80 ikke. Så kan de fortolke værdierne, og hvis jeg tager den her igen nu og simulerer op ad, så vil du se værdierne blive bedre og det er simulering af værdien i blodet. Og det kan man gøre på hver og en af dem her ikke, det her er bare ekspertorisk ilt så det ikke så vigtigt, den her er top tryk og det her er bund tryk, så man kan se hvor meget er tryk distentionen man også giver. Hernede er den simulerede værdi af pH i blodet, så hvis man kan se den er stor så kan man skrue ned for volumen, hvor sur bliver blodet, og hvor hurtigt bliver det surt. Det er så næste mulighed for dem der synes de skal have mere ekspertise, så nu kan de forstå hvad der skal ske med den her patient ikke. Det næste niveau er den her, og den forklarer fysiologi. Den her patient har 38 % af blodet som løber forbi lungenen uden at blive iltet. Der er aldrig nogen der kigger på den her side, det er noget man viser hvis man har en ekspert man skal snakke med systemet om. Så kommer man derind ikke, fordi de gerne vil vide det. Men det er måske øh, 10 personer i hele landet eller sådan noget.

39:18 Mikkel: En ekspert, det ville være en, altså.

39:23 IP: Det kaldes key opinion leaders ikke. Hvert land, for medicin studerende har man de 10 top personer som alle lytter til. Og de ved deres fysiologi, de ved hvordan patienten skal se ud. Så hvis man har en patient og man står der ved sengesiden, og man vil gerne snakke mere om patienten, så vil de sige "Hvordan ved jeg at patienten er tunet rigtigt, og er karakteriseret rigtigt?", så kan man gå herind og se det, de her ilt dissocationskurver osv. Så skaber man en tryghed i systemets funktionalitet, ved dem der har stor indflydelse til andre. Så det er det. Det næste man gør er at skrue på knapperne, systemet skruer ikke på knapper. Så her kan man lave de her simuleringer. Den anden vej *Swipe*, for forhistorie, og den får man med kurver, så med tiden går kan man se hvad rådet var, og er den blevet mere i centrum ikke. Hvis man trykker dobbelt på den, så kan man se fysiologisk information, og hvis man trykker igen så kan man se hvad rådet var og hvad de andre ting var ikke. Det gør jeg ofte, når jeg kommer ind på en afdeling hvor de bruger det så går jeg ind og ser hvad er der sket de sidste dage. Hvordan har rådet været, har folk fulgt med, er patienten blevet bedre. Det kan man følge med i, og hvis man trykker på den der, så får man lidt mere traditionel hvordan de XX 40:57 over tid, hvornår

kom blodgas ind og hvad var den værdi ikke. Så man få et overblik over, så det giver, vi håber, for jeg vil ikke sige det giver, men vi håber sådan bruges i ward rounds ikke. Når afdelingen organisere det to gange om dagen, og de går rundt til patienten ikke. Så kan man se hvordan patienten har det, og hvordan det er modificeret over tid, så kan man få et overblik på det ikke. Det er også en uddannelsesprocess der skal komme igang for at få det til at fungere ikke.

41:30 Mikkel: Jamen jeg tror faktisk det var det hele, er der noget vi ikke har været igennem?

41:35 Morten: Hvor ofte kommer med de der råd?

41:39 IP: Meget meget varierende ikke. Fordi hvis den optimisering regner ud at vi er langt væk, så, lad os sige at, at vi er på 700 og vi vil gerne have 500. Vi giver kun råd om 50 ml. Så den vil sige 50, og så ti minutter senere vil den sige 50, fordi vi venter på ligevægt igen. Så 50, 50 50, og såå, halv time, til en time senere så er man færdig med den løb der, og systemet synes det er optimal. Så vil systemet fortsætte med at regne, er der noget der er bedre? Er der noget der er bedre? Men patienten ændrer sig ikke, så kommer der ikke noget, fordi vi ha rnoget der er godt for patienten og patienten ændrer sig ikke. Så typisk vil patienten ligge der i nogle timer eller dage, hvem ved hvor langt, indtil de bliver mere rask, fordi det er det man gør i intense afdelinger, man kører på tid fordi de bliver raske selv. Og på et eller andet tidspunkt ikke, så begynder funktioner at forbedre sig og så kan man skrue ned for noget mere. Det vil systemet genkende, den vil sige "Nu kan jeg ikke forstå patienten mere, jeg vil gerne have noget blod". Så beder den om noget blodgas, og så indtaster man det, så bliver modellen tunet igen og så ændres noget mere. Og så på et eller andet tidspunkt vil patienten begynde at trække vejret selv, hvis man skruer ned for sederingen. De behøver ikke mere, de har ikke så meget smerte mere, så skruer vi ned for sederingen. Så begynder de at trække vejret selv, så kan systemet ikke genkende dem. Så beder den om blodgas, så foreslår den at skifte til en af den andre modes, hvor man kun giver tryk, da de trækker vejret selv, så skal man ikke bestemme frekvensen mere. Og så begynder den at skrue ned for det. Og så får man helt mangement flow i patienten på den måde ikke. Men det kommer i bursts ikke, massere af råd, intet, intet, intet. 5 timer, 6 timer, 2 dage. Patienten ændrer sig, så kommer der råd igen. 43:38 Morten: Så de vil se det når de går stuegang, eller hvad det hedder? Du sagde de kom to gange om dagen.

43:43 IP: Ja det er for at review patienter ikke, og lave planer for dem, men forhåbentlig så ser sygeplejeske det ved sengesiden. Intense afdelinger har en sygeplejeske per patient, typisk ikke. Nogle gange variere det fra land til land om det er 1 til 2 senge, nogle gange 1 til 3, men deromkring. Så der er konstant monitorering. 44:08 Morten: Ja så de er der hele tiden.

44:08 IP: Ja, og selvfølgelig har firmaet udviklet en ipad app, så man kan view alle patienter på samme tid. Så man kan sidde et andet sted og der kommer et råd og så gå hen til patienten. Men jeg tror kun det er implementeret i et sygehus nu, så den er endnu ny. Du kan jo se her, at den synes den op til 10 var en god ide, så nu vil den ændre mere, og fordi den kan forbedre ilt på den måde så kan vi skrue ned. Men den vil kun skrue ned med 5 fordi 5 er vores maks at skrue ned, så der er mange steps der kan komme der med at skrue ned.

44:51 Morten: Ja okay, så den vil ofte komme og sige.

44:55 IP: Den vil ofte komme og sige at den skal skrue ned på ilt indtil, men hvis man kender systemet så kan man gå ind her og se, at det ser fint ud længere nede, så lægerne kan være mere aggressive selv, men systemet vil ikke foreslog det. Systemet vil tage det i skridt. og hvis lægerne er aggressive selv, så lærer vi fra det, selvfølgelig ikke.

B.3 Analysis

This section contains the analysis of interview II. The analysis is seen in [Table B.2](#).

Table B.2: Analysis of interview II concerning the algorithms in the healthcare domain

Interviewsekvens	Meningskondensering	Analyse og fortolkning
00:06-04:42	<p>Post doctorate i '97'. En del af forskningsgruppen "Modelbaseret medicinsk støtte".</p> <ul style="list-style-type: none"> - Egen undergruppe "Respiratory and critical care group". - Fokus på beslutningsstøttesystemer for kliniske problemer relateret til lunge funktion og dybt sederet patienter. - Fokus på de aller sygeste patienter på sygehuset, som er kardiovaskulært støttede og støttede i cirkulatorisk aktivitet. - Tung viden om matematik beskrivelse af fysiologi. - Modellering af systemer ved hjælp af viden om blodet og dets transport af ilt+forskellige egenskaber. (kemisk modellering) - Modellering af elastens og eftergivenhed (modstand). (Mekanisk modellering) - Modeller kan benyttes diagnostisk, hvis de tunes til den individuelle patients data. - Modeller kan tune til specifikke patienter baseret på forskellige parametre, og derved hjælpe med diagnosticering. - 	<p>Personen har en PhD indenfor sundhedsteknologi. Har været en del af forskningsgruppen "Modelbaseret medicinsk støtte", og har nu egen undergruppe "Respiratory and critical care group".</p> <ul style="list-style-type: none"> - Har fokus på beslutningsstøttesystemer relateret til lungefunktion. Inkluderer dybt sederede patienter, som skal være kardiovaskulært støttede og støttede i cirkulatorisk aktivitet. - Personen besidder viden om matematisk beskrivelse af fysiologi. Samt modellering af systemer ved hjælp af fysiologisk viden, der kan tunes til specifikke patienter og hjælpe med diagnosticering. - Personen har været en del af udviklingen af Beacon Care system. Er bygget op omkring principperne om modellering af fysiologi. - Systemet (Beacon Care System) giver råd til sundheds- personale. Personalet skal selv udføre handlingen. Systemet prøver at lære af responsen.
00:06-04:42 (fortsat)	<p>Derved kan simuleringer af forløbet laves ud fra de modeller. Der viser hvilke effekter/konsekvenser forskellige handlinger har på patienten og patientens respons.</p> <ul style="list-style-type: none"> - De forskellige parametre har både positive og negative effekter, og disse kan så opvejes og balanceres. - Beacon Care system er bygget op omkring de principper bare matematik modellering af fysiologi. - Systemet giver råd til en læge/sygeplejerske og herefter er det op til mennesket at udføre selve handlingen. Systemet ser på responsen fra patienten og lærer af den. 	

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
05:16-15:13	<p>Firmaet "Mermaid Care" er et spin-off baseret på forskningen lavet. Her sælges Beacon Care system.</p> <ul style="list-style-type: none"> - Stephen er medstifter, og firmaet leverer nu penge videre forskning. - En tablet er udviklet til at ligge ovenpå respiratoren som brugergrænseflade. - I forskningsstudier for at teste systemet, har man fået gode resultater på forbedring af patienters tilstand. - Systemet er nu blevet CE mærket så det kan anvendes i Europa. - CE mærket giver kreditering og viser at systemet er sikkert at bruge. - 	<p>Beslutningsstøttesystemet Beacon er en kombination af mekanisk og kemisk modellering. Dette benyttes til simulering af ændringer i sygdomsforløbet, kommer med råd og kan måle disse råd. - Systemet tager ikke selv nogle beslutninger, men fungerer som rådgiver (open loop).</p> <p>-</p> <p>Systemet har været gennem forskningsstudier, der har til formål at teste og CE mærke systemet. Studierne har vist gode resultater og systemet er CE mærket.</p> <p>-</p> <p>CE mærket giver kreditering. Systemet er sikkert at bruge. Dette er en vigtig brik i den daglige brug af systemet. Uden det er det ikke muligt.</p> <p>-</p>

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
05:16–15:13 (fortsat)	<p>Nuværende stadie er at systemet er i gang med randomiseret test i Europa, med et mål af 700 patienter der skal anvende systemet.</p> <ul style="list-style-type: none"> - Istedet for enkelte læger der skal forstå systemet, skal der nu meget personale, heriblandt sygeplejersker, til at lære om systemet. - Afhængig af afdeling kan det være op til 200 mennesker (Intensiv fx.) der skal trænes i brugen af systemet - Firmaet kan betale sygehøuset for at ansætte forskningssygeplejersker, og træner også "Superusers". - Superusers er personale der går rundt og lærer andre brugen af systemet og sørger for at det fungerer. - Her kommer der mere en dagligdags interaktion af systemet istedet for forskningsforsøg hvor folk er "venlige" mod systemet. - TREAT systemet, anvendes indenfor antibiotika medicinering, og Rikke Pilmark har et firma der rådgiver om ernæring og glukose. - Indenfor respirator findes der to andre systemet, som er closed-loop. - Closed-loop betyder at systemet selv implementerer rådet på respiratoren uden menneskelig interaktion. - En stor debat er hvor gennemsigtig og forståelige et råd skal være. 	<p>Nyt stort studie. Randomiseret test i hele Europa, 700 testpersoner skal igennem med systemet.</p> <p>- Forskelse på forskningsstudier vs. randomiseret test. Enkelte læger skulle anvende det gennem forskningsstudier, hvorimod meget personale (Læger, sygeplejere) skal anvende systemet nu i randomiseret test.</p> <p>- Dagligdagsbrug af systemet. Superusers er trænet til at lære andet personale hvordan systemet skal anvendes.</p> <p>- Interaktion i daglig praksis er anderledes. Istedet for forskningslæger som er "Venlige" mod systemet, kommer det nu ud i daglig praksis hvor mange forskellige mennesker findes, og hvor interaktion med systemet skal fungerer for alle.</p> <p>- TREAT systemet. Rådgivning om antibiotika til læger.</p> <p>- Closed-loop respiratorsystemer. Implementerer rådet selv, uden at vente på personalets handling. Har det en effekt på hvorvidt man stoler på systemet?</p> <p>- Stor debat om transparens. Hvor gennemsigtig og transparent skal råd/systemer være?</p> <p>- Machine Learning. Black box modeller laves inden for machine learning, men der kan ikke sige hvorfor modellen agere som den gør.</p>

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
05:16-15:13 (fortsat)	<p>Indenfor AI er machine learning den mest prominente bevægelse. Her laves black box modeller, hvor problemet er at man ikke kan sige hvorfor den agere som den gør.</p> <ul style="list-style-type: none"> - Ved kontrolsystemer er det fint da man ikke behøves at forstå hvad der foregår. - Her prøver de(mermaid care) at få folk til at forstå hvorfor et specifik råd kommer. Dermed at gøre systemet og rådet mere transparent. - De (Mermaid care) støtter den tilgang der hedder at give personalet mere ekspertise ved at uddanne dem ved patientens sengeside, fremfor at systemerne bare gør det af sig selv. - Folk har forskellige meninger og fordomme, det er op til Mermaid Care/-Forskningen at bevise at det fungerer, for at få støtte. - Ting som reducering af ventetid på sygehus, ventetid på respirator, reducering af dødelighed er vigtige parametre. - Personalet skal være ligeglads med teknologien. Læger skal ikke interessere sig for AI, men om patienter. Og her skal de mødes med noget som kan hjælpe deres patienter eller deres process. - Hvis systemet kan hjælpe læger og deres patienter så skal de interessere sig, hvis ikke så skal de ikke. 	<p>Ekspertise frem for automation. Beacon Care System prøver at være transparent, og lære personalet om rådet og hvorfor det kommer. De prøver at fremme ekspertisen i personalet. Giver denne fremgangsmåde mere tro/tillid til systemet? Frem for de systemer hvor automationen er prioriteret?</p> <p>Forskellige holdninger, modbevist af forskningen. Et vigtig område der skal lykkedes for at få støtte er resultaterne/forskningen. Dette kan modbevise de fordomme der må ligge. Samtidigt vil reducering af ventetid og dødelighed hjælpe på administrativ/politisk støtte.</p> <p>Læger skal interessere sig for patienterne. Teknologien skal være ligevalgt, så længe det kan hjælpe patienter. Her skal læger mødes på halvvejen.</p> <p>Diskussionen skal ikke være om teknologi, men om patienter. Hvis teknologi kan hjælpe, så skal det hjælpe. Men der skal snakkes om hvordan systemer og råd kan gøres transparente så tillid i teknologien kan opbygges.</p> <p>Afgrænsning af råd. Systemet indeholder en funktion der kan afgrænse inden for hvilke intervaller råd skal forekomme. Her kan en læge give ansvaret videre inden for de fastsatte grænser. Er det sammen med tidlige nævnte funktioner, noget der er med til at give personalet en følelse af kontrol over det? Og dermed en større tiltro til det?</p> <p>Systemet vil stadigvæk kunne give råd udenfor grænsen men her kan en læge med ansvar så kontaktes</p>

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
05:16–15:13 (fortsat)	<p>I stedet skal der snakkedes om specifikke patient situationer, og hvordan råd og systemet kan gøres transparent så tro på systemet kan opbygges.</p> <ul style="list-style-type: none"> - Der eksisterer en funktion i systemet, hvor man kan sætte grænser for hvornår man vil have råd. fx. råd til denne patient hvis volumen holder sig inden for 400ml/pr. vejrtækning til 500ml/pr. vejrtækning. - Hvis volumen kommer udenfor grænserne vil systemet nævne at den stadig godt kan give råd men at det ligger udenfor de fastlagte grænser. Her kan en læge så kontaktes igen for at bestemme næste tilgang. - Derved kan en læge give ansvaret videre til sygeplejersker med støtte fra systemet. 	<ul style="list-style-type: none"> - Closed-loop systemer tager beslutninger selv. Det gør Beacon ikke. Closed loop systemer: SmartCare af Draeger og INTELLIVENT – ASV. Baseret på samme principper som Beacon Care, men tager beslutninger selv, derfor mere reguleringsystemer end beslutningsstøtte.
15:45–25:30	<p>Systemet tager aldrig selv beslutninger. Dette gør kun closed-loop systemer.</p> <ul style="list-style-type: none"> - Closed-loop systemer: SmartCare af firmaet Draeger. Et andet hedder INTELLIVENT – ASV. - Disse er mere reguleringsystemer end decideret beslutningstøtte, dog er de baseret på samme principper og fysiologi og modeller. - Systemet (MermaidCare) giver råd, men bestemmer aldrig. Derfor kan personalet ignorere alt som systemet fortæller dem. - Ingen FDA godkendelse i USA, til systemet, men andre closed-loop systemer har allerede godkendelse så det kan være muligt. 	<p>Beacon Care systemet bestemmer aldrig. Personalet står altid for at tage beslutningen, og kan ignorere systemet hvis de vil. Ansvaret for patienten ligger 100% hos personalet stadigvæk. Derfor ingen ansvar til systemet.</p> <ul style="list-style-type: none"> - Closed-loop systemer. Hvem har ansvaret i disse situationer? Closed-loop systemet er godkendt i USA, hvor de er meget 'cost' fokuseret. Hvis respiratoren arbejder selv, og får patienter hurtigere igennem er det en god ide. -

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
15:45-25:30 (fortsat)	<ul style="list-style-type: none"> - USA er meget kost fokuseret, og patienten skal gerne være hurtigt ude igen. - Der findes nogle overvågningsaspekter som en del af systemet. Man kan se alle målinger og hvilke råd der har været, hvilket betyder at man kan beskrive alle tidligere situationer patienten har været i. - Stephen kan gå igennem de seneste 14 dage for en patient, for at se om rådene er blevet implementeret af personalet eller ikke. - Der findes god overvågning der kan bruges til at forbedre systemet og kvaliteten og så dårlig overvågning der hjælpe med at "Slå folk på hovedet". - Teknologien er neutral men selve anvendelsen kan være både positiv og negativ men er også meget politisk præget. - I forskningsforsøg skal man have samtykke fra familie eller en læge med ansvar for patienten. Systemet er dog CE godkendt nu i Europa hvilket betyder at systemet findes i almindelig praksis. Som en del af lægens værkstøjskasse. - Positive/negative holdninger kan være på koncept, dvs. folk kan enten lide teknologi eller ikke lide det. - Holdning kan også være baseret på funktion, hvor nogen synes selve rådet er dårligt eller godt, afhængig er situationen. - 	<p>Overvågning igennem systemet. Som alt andet teknologi findes der overvågningsaspekter. Der er mulighed for at gå gennem data for en patient for de seneste 14 dage. Dette kan muligvis have en effekt på hvorvidt om systemet vil blive anvendt?</p> <p>- Samtykke fra patienter/familie. Systemet brugte i forskningsstudierne samtykke fra familier eller læger med ansvar for at teste systemet, da patienterne typisk er dybt sederet.</p> <p>- CE mærket i Europa. Da systemet nu er CE mærket er det at finde på sygehuse i Europa og også en almindelig praksis.</p> <p>- Forskellige typer holdning til teknologien/systemet. Positive/negative holdninger kan være på koncept, funktion eller selve anvendelsen</p> <p>- Koncept: Folk kan enten lide teknologi eller ej, hvilket ikke er velfunderede meninger og der blev nævnt tidligere at diskussion skal ændres fra det, til i stedet at handle om patienternes velfærd.</p> <p>- Funktion: Kan være på hvorvidt man mener det råd systemet afgiver er forkert/dårligt eller godt. Er meget afhængig af den specifikke situation.</p> <p>- Anvendelse: Små irritationsmomenter i anvendelsen af systemet kan ændre på opfattelsen af det. Hvis den ændrer på arbejdsprocessen eller forstyrrer vil det have negativ indvirkning.</p>

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
15:45-25:30 (fortsat)	<p>Holdning kan også være baseret på selve anvendelse, hvor irritation såsom at systemet beder om mere information eller kommer med råd/notifikationer på forkerte tidspunkter.</p> <ul style="list-style-type: none"> - Tidligere versioner viste en del irritation, hvor information om blodgas bl.a. skete tit. Hvis der kommer for meget irritation så giver folk op. - Holdninger på koncept/princip er lidt irrelevante da de ikke er velfunderede meninger. - Istedet for de meninger skal der snakkes om patienten og processen. Hvad var rådet, hvad skete der hos patienten, hvorfor var det godt/skidt. - Man skal huske at snakke om de dårlige ting for at kunne forbedre systemet. - En specifik situation hvor systemet har skabt irritation var ved flytningen fra seng til stol. Patienten bliver ustabil under flytningen, men stabil igen når han/hun er i stolen, og derved kommer der et nyt råd. Det skabte irritation da den ikke skal komme med nogen råd under sådan en manøvre. - Her kan en pauseknap anvendes til at sørge for at ingen råd bliver præsenteret og forstyrret. - Selv forskningen er overstået, nu er det integrationen med daglig praksis der er i fokus. <p>Vi præsenterer et råd, men ikke en forklaring. Hvis man vil have det, så skal man grave. Man skal grave i forklaringen til det niveau man vil have det.</p>	<p>En specifik situation med irritation. Systemet kom med råd på forstyrrende tidspunkter, hvilket irriterede personalet. Der skal sørges for at systemet passer ind i dagligdagen, og at interaktion med systemet ikke ændrer på processen.</p> <ul style="list-style-type: none"> - Systemet er væk fra forskningsområdet, nu skal det implementeres. Den svære del er at implementere systemet hos det almene sundhedspersonale. - Systemet præsenterer et råd, men ikke en forklaring. For at lære mere om selve rådet og hvorfor det kommer, er det muligt at grave i systemet.
15:45-25:30 (fortsat)		

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
26:27-44:55	<p>Systemet bliver vist på en tablet, og en gennemgang af interfacen går igang.</p> <ul style="list-style-type: none"> - <p>Systemet bliver vist i en demo version, med nogle simuleringer af målinger.</p> <ul style="list-style-type: none"> - <p>Stabilitet skal have været der i minutter ad gangen før systemet kan give råd. Når to minutter er gået vil den derfor begynde at regne på målingerne.</p> <ul style="list-style-type: none"> - <p>Blodgas giver blodets status, da systemet ikke kan udregne alt om kroppen.</p> <ul style="list-style-type: none"> - <p>Når systemet har udregnet gennem modellerne, så vil den lave en simulering og optimering.</p> <ul style="list-style-type: none"> - <p>Simuleringen viser hvordan patienten er lige nu, og hvad responsen vil være på ændring af ting.</p> <ul style="list-style-type: none"> - <p>Optimeringen er det optimale status for patienten nu, og hvis denne er langt væk, vil systemet give råd i den retning i steps.</p> <ul style="list-style-type: none"> - <p>Systemet kan godt tage lidt tid om udregning, men det betyder intet for patienten i praksis så der er massere af computonel tid.</p> <ul style="list-style-type: none"> - <p>Det er ikke bare for lægerne at undersøgelser laves på systemet, men ligeså meget administrationen som skal købe systemet. Der skal gerne være en økonomisk gevinst.</p> <ul style="list-style-type: none"> - 	<p>Simuleringen viser hvordan patientens status er lige nu. Samt hvad responsen vil være hvis noget ændres.</p> <ul style="list-style-type: none"> - <p>Systemet bruger tid på udregning. Da patienten er dybt sederet er der massere af computonel tid. Dette er måske ikke altid situationen for alle beslutningsstøttesystemer.</p> <ul style="list-style-type: none"> - <p>Der skal være en økonomisk gevinst. Det er bare at få læger til at anvende systemet at man laver undersøgelser, det er lige så meget for administrationen der skal købe det. Økonomien spiller altid ind.</p> <ul style="list-style-type: none"> - <p>Interfacen viser hvordan parametrene skal balanceres. Et spider-chart med hvilke parametre skal balanceres og holdes uden spikes så meget som muligt.</p> <ul style="list-style-type: none"> - <p>Der findes forskellige niveauer af information omkring patienten. Systemet er inddelt så der kan findes forskellig mængder af information hvis der er brug for mere viden om den specifikke patients status. Dette er tiltænkt forskellige personale.</p> <ul style="list-style-type: none"> - <p>Information skaber tryghed i systemet. Muligheden for at få information i dybden skaber en tryghed i systemets funktionalitet og brug. Bl.a. fra key opinion leaders.</p> <ul style="list-style-type: none"> -

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
26:27-44:55 (fortsat)	<p>Interfacen viser et spiderchart med forskellige parametre der kan ændres på.</p> <p>– Her handler det om at skabe en god balance, alle parametrene har positive og negative effekter, og de skal gerne balanceres så godt som muligt.</p> <p>– Interfacen viser et grønt område, hvor patienten gerne skal ligge men dette er ikke altid muligt fordi de ikke er raske.</p> <p>– Det skal helst undgås at have nogen spikes i ens spider chart men prøve at have en balance.</p> <p>– Demo versionen af systemet viser en patient der vil være dybt sederet (daglig tale koma), som ikke har spontan respiratorisk aktivitet og det er også her systemet anvendes i daglig praksis.</p> <p>– Sygeplejerskerne plejer at forstå det interfacen med spider chart og balancering af de forskellige parametre.</p> <p>– Lægerne vil gerne vide mere, og hvad det betyder. Derfor findes et niveau mere hvor simuleringen kan bruges af lægen og de kan se på forskellige værdier og hvad patientens respons vil være.</p> <p>– Det sidste niveau af information er tiltænkt key opinion leaders.</p> <p>– Det er information om fysiologien hos patienten, så de kan vide om patienten er blevet tunet rigtigt med de rigtige karakteristika.</p> <p>– Dette skaber tryghed i systemets funktionalitet og brug.</p>	<p>Fysiologisk historik. Patientens status kan over tid kan aflæses gennem systemet. Dette kan være en funktion der kan anvendes ved stuegang, som også er mål fra firmaets side.</p> <p>–</p> <p>Systemet prøver at optimere patienter i steps. Systemet er bygget efter at være 'forsigtigt'. Men som nævnt kan personalet manuelt ændre på indstillingerne mere aggressivt hvis dette ønskes. Herved lærer systemet også af det, til fremtidige patienter.</p> <p>–</p> <p>Systemet anvendes på intensiv. Systemet bliver brugt på intensiv og derfor er der ikke problemer med at selve rådet fra systemet bliver overset. Samtidigt er der udviklet en applikation til at skabe et overblik.</p>

Continued on next page

Table B.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
26:27-44:55 (fortsat)	<p>Key opinion leaders er ca. 10 personer som alle lytter til inden for det specifikke land. De kender til fysiologien.</p> <p>-</p> <p>Systemet kan også vise forhistorie, med hvilke råd er blevet givet samt fysiologisk information.</p> <p>-</p> <p>Stephen tjekker ofte dette på afdelinger for at se om rådet blev anvendt og om patienten er blevet bedre.</p> <p>-</p> <p>Håber på anvendelsen af denne funktion i stuegang, da man kan se om patienten er blevet forbedret over tid.</p> <p>-</p> <p>Råd fra systemet kommer meget varirende. Systemet går i steps mod hvad den optimale tilstand er for en patient, så derfor kan der komme mange steps/mange råd indtil at patienten har nået tilstanden. Hvis patienten ikke ændrer sig så kommer der ikke råd. Patienten kan derfor ligge i dage før noget nyt sker. Men når noget nyt sker vil systemet begynde at regne igen og herefter komme med nye råd til hvad der kan gøres for patienten.</p> <p>-</p> <p>På intensive afdelinger er der konstant monitorering så personalet ser nok rådene før de skal på stuegang.</p> <p>-</p> <p>Firmaet (Mermaid Care) har udviklet en applikation så alle patienter kan ses på samme tid, og samtidigt kan man se hvis der sker nogle ændringer/kommer nogle nye råd ved nogle patienter.</p> <p>-</p> <p>Selvom systemet kører i steps med råd, så vil lægerne altid kunne ignorere dem og være mere aggressive i deres ændring af respirator indstillingen. Systemet vil prøve at lære af de aggressive ændringer.</p>	
26:27-44:55 (fortsat)		

C

Interview III

This appendix contains the interview guide, the transcription, and the analysis of interview III.

C.1 Interview Guide

This section contains the interview guide for the third expert interview conducted during this study. The interview guide can be seen in [Table C.1](#).

Table C.1: Interview guide III

Introduktion: Vores projekt omhandler tillid til algoritmer, heriblandt beslutningsstøttesystemer. Vi har i vores projekt valgt at fokusere på sundhedsdomænet.

Vi er igang med en eksplorativ undersøgelse, der undersøger, hvordan sundhedsprofessionelle opfatter råd, både generelt og fra IT-systemer. Det endelige mål med vores undersøgelse er at kortlægge, hvilke opfatninger der påvirker vægtning af råd, både positivt og negativt, og hvilke opfattelser der ikke gør.

Den eksplorative undersøgelse ligger til grund for en mere systematisk undersøgelse, hvor en række forsøgspersoner skal tage stilling til nogle udsagn og vurdere en række 'rådgivere' baseret på disse. Det er derfor vigtigt for os at finde nogle udsagn, som kan danne grundlag for denne mere systematiske undersøgelse.

Anonymitet: Dit navn anonymiseres, hvis dette ønskes.

Båndoptagelse: Interviewet optages, transskriberes og benyttes i forbindelse med vores projektrapport. Ønskes dele gennemlæst kan dette arrangeres.

Tid: Interviewet vil tage maks en time.

Forskningssspørgsmål	Interviewspørgsmål
Hvem er interviewpersonen?	<ul style="list-style-type: none"> • Kan du fortælle lidt om dig selv? <ul style="list-style-type: none"> - Hvad er din baggrund? - Hvad er din rolle her hos Region Nordjylland?
Vi vil høre om beslutningsstøtteprojektet. Hvilken type beslutningsstøtte og hvordan fungerer det? Har interviewpersonen tidligere erfaring med beslutningsstøttesystemer? Er der erfaring med at implementere andre/lignende beslutningsstøttesystemer?	<ul style="list-style-type: none"> • Vi vil gerne høre lidt om det regionale beslutningsstøtteprojekt, kan du fortælle os lidt om, hvad det handler om? • Hvilk type beslutningsstøtte er der tale om i projektet? • Har du tidligere været involveret i implementeringen af beslutningsstøttesystemer? • Hvilke systemer/projekter var dette?
Hvilke bekymringer har der været blandt klinikere om beslutningsstøttesystemer? Er der nogle generelle holdninger til beslutningsstøtte blandt klinikere? Er der forskellige lejer blandt klinikere med hensyn til beslutningsstøtte? Har interviewpersonen eksempler på vendinger, situationer eller lignende, hvor klinikere udtrykker deres specifikke holdninger?	<ul style="list-style-type: none"> • Hvordan er den generelle holdning til beslutningsstøttesystemer hos klinikere? <ul style="list-style-type: none"> - Findes der forskellige lejer? • Har der været bekymringer relateret til det nye beslutningsstøtteprojekt? • Hvilke bekymringer har der været og hvorfor tror du at disse findes? • Kan du nævne nogle specifikke situationer eller bekymringer som er kommet til udtryk i forhold til klinikeres holdning, og hvad har deres grundelse været for bekymring?
Er det muligt at holde nogle korte interviewsessions (ca. 30 min.) med nogle klinikere for at få noget data fra dem direkte? Kan det nås inden påske? Har de tid i påsken? Kan det ske umiddelbart efter påske?	<ul style="list-style-type: none"> • Er det muligt at holde nogle korte interviewsessions (ca. 30 min.) med nogle klinikere for at få noget data fra dem direkte? <ul style="list-style-type: none"> - Kan det nås inden påske? - Har de tid i løbet af påske? - Kan det ske umiddelbart efter påske?

C.2 Transcription

This section contains the transcription of interview III. The transcription was made using the oTranscribe website ([Bentley, 2019](#)).

Transcription

00:00 Introduktion til vores projekt og hvad vi prøver at finde ud af.

06:12 Mikkel: Jamen vi kan bare starte med, sådan lige at fortælle lidt om dig i forhold til hvad din baggrund er.

06:23 IP: Jeg er uddannet sygeplejerske for mange mange år siden, øhh, har ikke arbejdetude i praktikken i mange år. Jeg har, puhh, jeg har lavet alt muligt, jeg har undervist og jeg har pas.., jeg har arbejdet med døve børn, og alt muligt, og så fik jeg ansættelse her for, er det ni år siden i IT, hvor jeg lige skulle hjælpe med at implementere det fælles medicinkort, og så har jeg siden da været projektleder for det. Og det arbejder jeg stadigvæk med, vi har stadigvæk nogle hængepartier og nogle ændringsønsker og noget til det.

07:02 Mikkel: Ja.

07:02 IP: Og i forbindelse med det, da vi startede det fælles medicinkort, øhh, da vi implementerede det i 2012, var det lægernes ønsker at, at vi fik noget omkring, øhh, dosisdispensering og, øhm, antikoagulationsbehandling og noget beslutningsstøtte. Og derfor blev det i 2015 aftalt at der skal laves beslutningsstøtte i alle systemer. Både lægepraksissystemer og alle regionale systemer, øhhm, og i den aftale var der to nye ting. Et at det var en region der skulle være projektledende, to at regionerne skulle betale for udviklingen af de praktiserende lægers systemer og det er fordi formålet er at nedsætte medicineringsfejl og undgå medicineringsfejl i sektorovergangen og undgå genindlæggelser på grund af medicineringsfejl, og hvis vi kan få nogen effekt ud af det så er det regionerne der er gevinstejere og derfor blev det pålagt at regionerne skulle betale.

08:57 Mikkel: Øhhmm, det her nye regionale beslutningsstøtteprojekt, var det det du var ved at fortælle om, medicinering osv.

09:05 IP: Ja.

09:05 Mikkel: Hvad er det præcis det handler om, er det medicinering hos.

09:17 IP: I den aftale der blev indgået, der indgik der tre ting. Der indgik, der skulle være maksdosering, og der skulle være interaktioner og der skulle være risikosituationslægemidler og så CAVE. Det var de punkter der indgik i økonomiaftalen med staten og staten siger: I skal lave det her. Og så fik vi projektet heroppe og så begyndte vi at kigge på det her. Det der så sker det er at jeg begynder at kigge på det her datagrundlag. CAVE den er sådan lidt for sig selv, fordi forskellen mellem det når man sådan snakker tekniske løsninger til det, så er CAVE noget som handler om: 'Hvad er det personen ikke kan tåle'. Og lige så snart det er personhenførbart så skal det sendes på CPR-nummer, så vi sagde egentligt bare: Aaarrhgh prøv lige at del den op i noget der er lidt mere komplekst og så noget der egentligt bare sendes, at her handler det jo ikke bare om at personen får di-coxin. Men dicoxsin i det hele taget, så her kan du sende oplysningerne uden at du har CPR-nummeret med. Og det gør en stor forskel i den lovgivning der er i forhold til, hvordan du laver dine tekniske løsninger. Så kigger vi på, hvad har vi i Danmark. Det fælles medicinkort det er drevet af sundhedsdatastyrelsen og det er faktisk første gang at staten driver et IT-system. I den process har de måtte lære meget og de har stadigvæk meget at lære. Aftalen var egentligt at sundhedsdatastyrelsen skulle levere selve maskinen og selve beslutningsstøtten.

12:01 Mikkel: Ja.

12:01 IP: Så gik jeg ind og kiggede på de her risikosituationslægemidler og dem kan i google, der er sådan en liste over risikosituationslægemidler, og de er meget, det er sådan meget generelle, sådan nogle som insulin og, altså sådan nogle hvor det kunne tænkes at du skal have beslutningsstøtte, for dem som får dem hele tiden. Men de er meget opslået, de står meget på den der liste på grund af administrationsfejl, altså når det er hjemmeplejen der skal give det her, for eksempel, hvis de skal give noget blodfortyndende så er det afhængigt af en blodprøve og så bliver blodprøven ikke taget eller også så får de det ikke eller, så det her det var meget sådan administrationsvejledende, så egentligt så sagde vi at det er så ring at det kan vi ikke bruge. Interaktioner, der har vi den danske interaktionsdatabase og den er, det er den de bruger i dag, de arbejder på at forbedre den, der er plads til forbedringer men den er kendt af alle. Og den har alle for.., og den har alle i dag. Så der var jo ikke noget nyt i der her ved at få beslutningsstøtte på interaktioner, når de fik det i forvejen fra den samme database. Så var der maksimumdoseringen og det sundhedsdatastyrelsen kunne levere i første omgang det var på, jeg kan ikke huske det men ca. 10 præparerer og kun for voksne og så egentligt den der med at det, hvis du har en tastefejl og kommer til at skrive 100 i stedet for 10 så kan du også få beslutningsstøtte på det. Og så sagde vi kig lige ned her og se i forhold til hvor mange millioner vi bruger på det her og regionerne skal betale, så sagde vi det her det er no-go. Det får vi, det tager minimum 10 år hvis Danmark skal opbygge et datagrundlag der er godt nok til at give beslutningsstøtte på så vi kan få en gevinst ud af det. Så vi sagde: Det her kan vi ikke bruge. Nå. Hvad gjorde vi så? Så sagde vi: Okay vi går i udbud, vi går ud i verden og ser, hvad har de

af datagrundlag derude. Og så gik vi i udbud, og i udbuddet der skulle vi købe et datagrundlag og et beslutningsstøttesystem, altså en teknisk løsning, som kunne gøre at vi har systemet her, at vi skal sende ind-data og så få et svar ud, altså ud-data. Her kan man sende ind: Alder, vægt, gfr-værdi, køn, graviditet, amning, alt muligt. Hvad har vi af data vi kan sende ind, og hvad får vi ud i den anden side, det handler jo meget om, altså selve systemet som sådan er jo meget logisk, send noget ind, noget ud, men det var det der var vores store udfordring det var jo datagrundlaget, hvordan får vi fat i noget der er ordentligt. Så gik vi i udbud i EU-udbud og der fik vi tre svar på den. Vi fik et tilbud med et amerikansk datagrundlag, og et med et finsk og så ind fra højre der kom der et med et dansk fra dansk lægemiddelindustrien, faktisk det der er lig med promedicin.

16:11 Mikkel: Ja.

16:11 IP: Og dem kiggede vi på og valgte faktisk et firma som kunne levere systemet med dem her som underleverandører. Firmaet her, Trifork, det er de samme som drifter det fælles medicinkort i dag. Og så får vi så beslutningsstøtte, altså fra ala, vi får også fra dli eller promedicinn bare, giver svaret og så havde vi jo kigget meget på det her og har haft en lang proces omkring de her ting og ligesom sagt: hvad kan vi forestille os? Og vi havde et dialog-præget udbud, forstået på den måde at for at de her og vi ligesom kunne finde ud af, hvad var det vi skulle ud og lave, så mødes vi lige nogle gange og snakker om, hvad kan i give og hvad tænker i, for at blive dygtigere i den proces, øhmm, så der hvor vi er lige nu, det er at vi har valgt Trifork med DLI som underleverandør og vi står lige for at skulle teste datagrundlaget, eller selve systemet fungerer sådanset udemærket men nu er vi jo så ude i det der datagrundlag og det er jo en sværere proces, men, og så er vi nødt til at kigge på om de der data vi skal kunne sende ind, om de kan håndtere det, nogen af de her data havde vi jo meget i selve processen også, om hvad kunne vi tænke os, jamen vi kunne tænke os beslutningstøtte i forhold til det og det og det og det. For eksempel sådan noget som leversygdomme det var vi nødt til at trække fra fordi vi egentlig ikke kunne finde nogen eller blive enige om, hvilke data vi skulle sende ind for at få noget ud igen. Lige nu har vi lidt en, ja en udfordring i forhold til diagnoser fordi datagrundlaget kan godt give diagnoser på indikationer, øhh, og systemet kan godt leve svaret, men vi har simpelthen ikke noget ordentligt at sende ind. Og det handler om den måde de arbejder på, altså heroppe arbejder, i Region Nordjylland arbejder vi med arbejdsdiagnoser og sygehulslægerne sætter faktisk først diagnoserne når de bliver udskrevet, øhh, de praktiserende læger de er sådan sjældent diagnoser når de bliver indlagt, men folk skriver ikke den der udskrivningsdiagnose ind et sted, hvor den ligesom kan tages fra. Så der har vi en, det ved jeg ikke, en stor opgave fremover at finde ud af hvordan vi i både regionerne og ved de praktiserende læger og på tværs af de der sektorer kan blive bedre til at have nogle data vi kan få noget mere kvalificeret beslutningsstøtte på. Alt det her, i det her, der er tre pilotprojekter, et for Region Nord og så er der to lægepraksissystemer som indtil Januar 2020 skal prøve at lave en løsning som kan vises i brugergrænsefladen. Og det er vi sådan så småt begyndt på. Hvis vi så lige hopper tilbage igen, øhmm, fordi det her det hedder beslutningsstøtte i det fælles medicinkort og i regionerne arbejder vi både på det fælles medicinkort og når patienten er indlagt i vores lokale system så vi er nødt til at skulle have beslutningsstøtten ind begge steder og det er vi så ved at se om vi kan implementere, der er jo stor forskel på sådan et lægepraksissystem som har x antal praktiserende læger og der sidder en udvikler der lige kan lave noget og så sende ud til tre læger og så sige, hvordan er det her frem for at vi kan, når vi går i pilot så er vi nødt til at skulle ud til alle i hele regionen. Så den opgave er meget større end de andre. Hvis vi så lige går tilbage og siger at det fælles medicinkort, det er sundhedsdatastyrelsen der drifter det og har gjort det fra starten. Så kom rigsrevisionen på et tidspunkt og sagde: hmm, hvorfor har staten udbetalt 200 millioner kroner og regionerne har ikke implementeret endnu? Hvad ... laver de? Så gik rigsrevisionen ud og så på sundhedsstyrelsen og så på sundhedsdatastyrelsen og på regionerne. Og så sagde de: Sundhedsstyrelsen hvad har i gang i hvorfor har i udbetalt de her penge uden at lave opfølgnings på, hvad i får for pengene, hmm, det kunne de ikke rigtigt svare på. Sundhedsdatastyrelsen, jamen hvordan styrer i det her? Har i en styregruppe og har i lavet risiko, der var mange ting de ikke havde og de fik de fik flest skæld ud og sundhedsstyrelsen fik nummer to og regionerne fik altså også noget skæld ud. Noget af det vi blev kritiseret for det var at der ikke var lavet, arbejdsgangsanalyser og brugerkrav. Og så egentlig at, spørgsmålstegn ved, eller noget kritik ved at implementeringen ikke havde været god nok. Men hvis lige vi prøver at sætte den væk, så gjorde vi det at da vi fik beslutningsstøtteprojektet heroppe, så tænkte jeg nu må vi hellere tage udgangspunkt i rigsrevisionens kritikpunkter og sige hvad fejl har vi gjort, og kan vi lære af dem? Så det første vi gjorde egentlig inden at vi kom så langt at vi gik i udbud, det var og gå ud og lave arbejdsgangsanalyser og brugerkrav og lave sådan noget interviewguide og var ude at spørge, gud og hver mand af læger og sige, og grund til at jeg siger det lige nu det er at det der materiale det må i godt se, det er selvfølgelig nogle år gammelt, men det siger noget om hvordan de så på det på det tidspunkt og hvad det var de godt

kunne tænke sig, øhmm, og så lavede vi sådan nogle sådan arbejdsgangsanalyser, det ved jeg ikke om vi fik så meget ud af, men i hvert fald de der brugerkrav, og der sendte vi, der gjorde vi det at vi heroppe, lavede i region nord og egentlig også sendte dem ud til de andre regioner, det fik vi ikke så meget ud af, men så betalte vi to lægeprakssissystemer om at gå ud og lave brugerkravene og den måde vi gjorde det på var lidt forskellige og det vi ligesom meldte tilbage, og det vi kunne konkludere ud af dem var også lidt forskellig, hvad det var man gerne ville have. Så vidt jeg kan huske sagde de praktiserende læger: At det de kunne få mest ud af umiddelbart var, heheh, ja det lyde helt åndssvagt, var beslutningsstøtte til gravide og ammende og så den der ældre, jeg tror ikke de, jeg kan ikke huske, nu skriver jeg besværlige medicinske patienter, altså den der gamle medicinske patient som får meget medicin. Hvordan finder vi ud af at fru Hansen er 89 og hun fejler alt det her og hun får alt det her medicin, hvordan kan jeg som praktiserende læge være sikker på at det alt det der 20 slags medicin hun får det interagere med hinanden og så havde de et stort ønske om de, altså de ammende og de gravide. Fordi hvis man sådan kigger på det så må de gravide faktisk ingenting få og heller ikke når de er ammende og hvordan finder man ud af det og og det har vi lidt døj med nu her og vi snakker om, hvad er det for nogle ind-data vi skal sende. Gravide er det alle mellem 13 og 63 eller og så kom vi jo til at snakke om her den anden dag, kan du være mand og være gravid, det havde de jo egentligt trukket fra og så blev vi enige om når, men principielt kan man jo godt. Hvis du får lavet kønsskifte, ikke. Altså, Ja.

26:42 IP: Så det var egentlig det de konkluderede, noget af det vi gerne vil have det var jo selvfølgelig også til de der komplekse medicineringer og så var der et stort ønske om noget beslutningsstøtte til børn og, der var sådan en lang liste med noget af det vi konkluderede, men det var lidt forskellig, hvad det var, man egentlig sådan gerne ville have. Der blev også spurgt lidt forskelligt, altså selve den måde interviewene, eller det der blev interviewet på var også forskelligt. Men det gav et forskelligt billede, specielt med den der. Og lige nu kigger vi også lidt på de der inddata med gravide, altså hvordan, skal du altid spørge dem: Er du gravid? Hmm, skal du så kunne sætte kryds for at få den relevante beslutningsstøtte eller, eller hvad og hvornår er du ammende, altså er det en periode efter du er gravid eller er det, hmm, eller skal man bare sige at alle altid er gravide og ammende og så, ja det er nogle af de der ting vi arbejder lidt på nu som er sådan lidt besværlige. Det er sådan set status lige nu, lige nu skal vi lige have testet her inden for det næste måned, to måneder alle de der inddata, hvor man ligesom siger at det vi har stillet krav til, at systemet skal kunne give svar på, kan det så også det. Øhhm, og vi har været ved at kigge lidt på det og i den forbindelse kommer vi jo til at kigge på, hvad er det så for nogle svar vi får ud. Hmm, for noget af det vi i hele den her proces har sagt fra starten det er: Alfa omega er at man har tillid til det svar der kommer, der er sat nogle meget vold, arrh jeg vil ikke sige voldsomme, nogle stramme krav til datagrundlaget, altså hvor kommer det fra, hvem fører det, hvem laver det hvilken organisering står bag. Det er jo vigtigt at det er videnskabeligt bevist, der en meget stærk videnskabelig organisering bag DLI, det der så har været meget kritik på det er jo så at Dansk Lægemiddelindustri er jo, er jo et privat firma, så. Og det har vi jo også sat krav til at, jamen kan vi være sikre på at de er uvildige. Det siger de så. Men, vi begynder jo sådan at kigge lidt. Hmpfh, jaa, datagrundlag, ja, er der, kan, har man tillid til svaret, og det er det vi har set en lille smule på og skal til at kigge noget mere på nu her.

30:09 Mikkel: Ja

30:10 IP: Sådan at vi når vi engang i efteråret eller lige først inde i det nye år implementerer det her, så skulle vi rigtig gerne have at, at de har tillid til systemet, eller at vi i det mindste kan besvare de spørgsmål de har. Hvorfor gør det sådan? Det er fordi at. Og det er det der giver tilliden, ja.

30:40 IP: Det var rigtig mange ord på kort tid.

30:42 Mikkel: Ja. Men det var.

30:46 Morten: Men det er jo også det vi synes der er rigtig interessant at undersøge, netop det med tillid og hvad det er der gør...

30:51 Mikkel: ...at man har tillid.

30:51 Morten: Ja.

30:55 Mikkel: I forhold til de her, øhh, undersøgelser der er blevet lavet så, øhhm, hvad har så, jeg ved ikke om det er noget der er blevet kigget på også, men hvad er sådan den generelle holdning. Hvordan har den generelle holdning været til det der med at få et system sådan her ud? Har der været nogle forskellige lejre? Eller?

31:15 IP: Noget, det de kan levere her, Trifork, det vi for leveret det er 11 moduler, der har selvfølgelig også været lavet nogle, altså 11 beslutningsstøttemoduler. Den ene er gravide og ammende og diagnoser og maksimum og gfr-værdi og øhh, antibiotikavejledning og nu kan jeg ikke huske flere, men det er også ligemeget, men ligesom 11 af de her. Noget det har systemerne i forvejen. Noget er et krav fra fmk, for eksempel at du får en advarsel mod dobbeltordinationer

fordi vi på et tidspunkt på fmk havde et enormt antal af dobbeltordinationer så sagde vi, hvordan får vi det bragt ned? Alle skal have en advarsel. Det kommer der også her og det du får beslutningsstøtte på her er et bedre datagrundlag end det der kan gives derovere så vi skal lige have fundet ud af at de skal jo ikke have samme advarsel to gange. Men hvordan håndterer vi så det. Der blev også i forhold til interaktionsdatabasen herovere besluttet den der klinikergruppe som var med til at lave de her krav, at vi tager det danske fordi der er faktisk en norsk der er bedre men ligeså snart du er ude i verden skal du have omkonverteret data.

33:08 Morten: Ja.

33:08 IP: For eksempel både for det finske og det amerikanske, det er nemt nok, hvis en panodil er 500 mg af et stof er lig med det her, men lige så snart du er ovre i kombinationspræparaterne, hvor du har et præparater der indeholder to ting, så kan du faktisk ikke, hvad hedder det over dem, og det gjorde faktisk at nogle af dem her fravalgte vi fordi at kombinationslægemidlerne faktisk udgjorde 15 % af dem, det vil sige at du fra starten ikke kan få beslutningsstøtte på alle præparater, men du kan tage de 15 % ud og så kan du tage alle ud vi i dag ikke har noget datagrundlag på for eksempel alle de der xx 33:57 præparater, altså dem der er lavet på privatapotekerne, dem har vi ikke noget officielt datagrundlag på, så det kan vi ikke men så satte vi krav om at, hvis du opretter en ordination du ikke kan få beslutningsstøtte på så skal du have at vide at det her det får du ikke beslutningsstøtte på. Så noget har de i dag og det er meget forskelligt. Når du kigger på nogle af de der krav fra de der spørgeskemaer som vi, eller fra de der konklusioner som vi lavede ud fra de der spørgeskemaundersøgelser så vil de gerne have støtte, men de gider ikke lægge det væk. Altså de vil ikke have en ekstra arbejdsgang, de vil gerne have hjælp til det men vi har slægt fast med syvtommersom at nu får i noget hjælp, men det er jeres ansvar, det er lægen der ordinerer ligegyldigt hvad beslutningsstøtten så end siger. Så han skulle gerne få hjælp til det men det er ham, det er hans ansvar, og de der gamle læger de kan godt, med mindre de er ovre i den der meget komplekse medicinering, hvor det bliver rigtigt vanskeligt at finde ud af hvad der er hvad her, eller hvor mange af de her præparater interagerer med hinanden.

35:33 Mikkel: Ja. Det var fordi vi kom faktisk også til at læse lidt omkring det her beslutningsstøtteprojekt, som vi lige kunne finde, at der stod lidt om at der også havde været lidt skepsis også. Hvilke, ligesom, bekymringer har der været omkring det nye her?

35:55 IP: Regionernes læger bliver jo ikke så tit spurgt, vel, her går regionerne ud og laver en aftale og staten siger i får de her penge, tak dem vil vi gerne have og så laver vi det her og så sender vi det ud, de bliver ikke sådan spurgt, men i den anden lejer, der står PLO, de praktiserende Lægers Organisation. Og det var noget trægt og få dem med ind i det her, arrgh, det tog nok et års tid. De vil gerne kunne slå det fra, men de gik faktisk ind og sagde at det var okay, de ville gerne gå med i processen, de har nogle krav til at det, og vil ikke bruge det, hvis de synes at det genererer for meget. Og det er forståeligt nok, men de skal jo selvfølgelig også have tillid til det. Men indtil videre ser det fint ud. Her i næste måned har vi sådan en workshop, hvor de der tre pilotsystemer skulle prøve, sådan deres første spæde mock-ups på, hvordan det her kunne komme til at se ud, at de prøver at vise det for den der nationale klinikergruppe, der er sådan en national klinikergruppe med klinikere fra hver region og repræsentanter fra PLO.

38:41 Mikkel: Så har der været nogle specifikke ting i forhold til det her med i forhold til, at du snakke om at, at de skal have tillid til at rådet er rigtigt eller at den kommer med et kvalificeret bud. Har der været nogle specifikke holdninger der har været udtrykt igennem undersøgelsene eller?

38:59 IP: Vi har jo diskuteret meget, hvordan finder du ud af det, altså, hvordan kan du finde ud af at du har tillid til det her. Og hvem kan teste det, du kan jo aldrig teste det fuldt ud, du kan bare lave stikprøver, og hvordan skal vi såøre det? Så derfor gjorde vi egentligt det at vi gik ned og sagde at de krav vi sætter i vores udbud til datagrundlaget er at for at finde noget. Og så satte vi nogle krav til ind data og ud data. Ja.*Begynder at finde nogle dokumenter til os om krav og deres undersøgelser*

40:20 Mikkel: Ja, men jeg tror egentlig også at det var de spørgsmål vi havde i forhold til projektet der, så vil vi gerne høre omkring den her mulighed for at må interviewe nogle klinikere.

C.3 Analysis

This section contains the analysis of interview II. The analysis is seen in [Table C.2](#).

Table C.2: Analysis of interview II concerning the algorithms in the healthcare domain

Interviewsekvens	Meningskondensering	Analyse og fortolkning
00:00-08:57	<p>Interview personen har tidligere erfaring indenfor sunheddomænet i form af sygeplejeskefaring og arbejde med døve børn.</p> <p>-</p> <p>Sidenhen har personen været ansat hos Region Nordjyllands IT afdeling ifht. implementering af nye systemer til sundhedssektoren.</p> <p>-</p> <p>Har været med i implementeringen af de fælles medicin kort.</p> <p>-</p> <p>Gennem implementeringen fik IT forespørgelser på dosisdispenseringer, antikoagulationsbehandling og beslutningsstøtte.</p> <p>-</p> <p>Projektet om beslutningstøtte blev aftalt i 2015, hvor det skulle implementeres i lægepraksissystemer og regionale systemer.</p> <p>-</p> <p>Formålet var at nedsætte medicineringsfejl, og undgå medicineringsfejl i sektorovergang, samt at undgå genindlæggelser på grund af disse fejl.</p> <p>-</p>	<p>Praktisk erfaring. IP har erfaring indenfor sundhedssektoren og blev ansat til at hjælp med implementering af nyt system FMK.</p> <p>-</p> <p>Erfaring fra tidlige projekter og nu på det nye IT projekt som leder. IT-leder på de nye beslutningsstøtteprojekt for regionerne, med pilottest i Region nordjylland.</p> <p>-</p> <p>Formålet med projektet er at ned-sætte fejl. Medicineringsfejl generelt og i sektorovergang, samt reducering af genindlæggelser.</p> <p>-</p>

Continued on next page

Table C.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
08:57-30:40	<p>Aftalen om beslutningsstøtte indeholdte fire ting/områder: Maksdosering, Interaktioner, Risikosituationslægemidler, CAVE</p> <ul style="list-style-type: none"> - Region Nordjylland fik projektet af staten. - CAVE er forskelligt fra de andre områder, da det handler om personfølsomme data. Her skal CPR nummer indover, og derfor blev CAVE opdelt fra de andre tre områder. - De tre andre områder indeholder information som ikke skal findes gennem CPR nummer. Dette gør en stor forskel i forhold til lovgivning og tekniske løsninger. - Herefter blev der set på hvilke andre tekniske systemer der er i Danmark. Det fælles medicinkort er første gang staten har drevet et IT-system. - I implementeringen af FMK har staten lært meget. - Risikosituationslægemidler fx. insulin. Her sker der tit administrationsfejl og derfor skal noget nyt udvikles til at administrere dem. - Der findes den danske interaktionsdatabase, som bliver anvendt idag og stadig bliver forbedret. Den anvendes af alle idag og derfor gav det ikke mening at udvikle noget nyt til dette område da det ville stamme fra samme database alligevel. 	<p>Beslutningsstøten kommer på forskellige områder.</p> <ul style="list-style-type: none"> - CAVE er personfølsomdata der skal sendes gennem CPR nummer. Dette var ikke i fokus for projektet da det er noget andet indenfor lovgivning og teknisk løsning. - Ikke personfølsomme data områder. Maksdosering, risikosituationslægemidler skal der udvikles beslutningsstøtte til. - Datagrundlag i EU udbud. Forskelige forslag kommer ind men Dansk Lægemiddel INdustri vælges som leverandør af datagrundlag. Trifork som systemsleverandør.

Continued on next page

Table C.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
08:57-30:40 (fortsat)	<p>Maksimumsdoseringen er når en tastefejl gør at der bliver skrevet fx. 100 istedet for 10. Her skal der være beslutningsstøtte på. Men der findes intet datagrundlag for at kunne lave denne støtte.</p> <ul style="list-style-type: none"> - Region går i EU udbud for at finde et datagrundlag samt et beslutningsstøttesystem - Systemet fungerer ved at data bliver sendt ind og vurderet på forskellige parametre og ud kommer beslutningsstøttende råd - Forskellige tilbud på datagrundlag kommer ind, heriblandt dansk lægemiddelindustri. - Firmaet valgt til at levele beslutningsstøtte er Trifork, som også drifter FMK. - DLI bliver valgt som leverandør af datagrundlag, men processen med at vælge hvilke data der skal anvendes til at give svar er svær. - De data der skal bruges til beslutningsstøtten er ikke fastlagt endnu, da forskellige udfordringer findes. - Bl.a. findes der ikke meget data på indlæggelser, og i region nordjylland bliver der arbejdet med arbejdsdiagnoser (Matcher Pernille Interview med den ekstra notesbog), og diagnoser bliver først skrevet ind når patienten bliver udskrevet. - Stor udfordring i at få data som beslutningsstøtten kan bruges på. Lægerne arbejdsprocess skal muligvis ændres? For at systemet kan bruges. - 	<p>“ind” data skal stadigvæk defineres, hvilket er en svær opgave.</p> <ul style="list-style-type: none"> - Arbejdsprocessen i Region Nordjylland er med arbejdsdiagnoser, hvor de reelt først bliver skrevet ind ved udskrivning. Dette gør at de data der skal bruges til at vurdere ordinatioen, ikke bliver skrevet ind før patienten er udskrevet igen. - Der skal måske ændringer til i arbejdsprocessen. - Region nordjylland skal ud til alle læger i regionen. - Brugerkrav og arbejdsgangsanalyser er vigtige elementer i udviklingen. Disse er kommet som krav ud af implementeringen af FMK. -

Continued on next page

Table C.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
08:57-30:40 (fortsat)	<p>Tre pilotprojekter findes, et i region nord og 2 i lægepraksissystemer som skal prøve at lave noget brugergrænseflade.</p> <p>-</p> <p>Der er stor forskel på implementeringen af systemerne, da regionen skal ud til alle i hele regionen mens de andre skal ud til 3-4 praksis læger.</p> <p>-</p> <p>Rigsrevisionen lavede en undersøgelse af sundhedsdatastyrelsen og her kom der kritik af implementeringen af FMK. Der mangede arbejdsgangsanalyser og brugerkrav, samt at implementeringen ikke havde været god nok.</p> <p>-</p> <p>I det nye beslutningsstøtteprojekt fik de herefter lavet arbejdsgangsanalyser og brugerkrav, gennem interviews og spørgeskemaer.</p> <p>-</p> <p>De to andre lægepraksisystemer blev også bedt om at lave undersøgelserne. Forskellige resultaterne fandtes.</p> <p>-</p> <p>Praktiserende læger kunne få mest ud af beslutningsstøtte til gravide og ammende, og ældre (Besværlige medicinske patienter)</p> <p>-</p> <p>Der findes mange parametre for at bestemmer de her data som skal bruges til systemet. Hvilke aldre? Hvornår kan man være gravid vs ikke gravid? Mand vs kvinde?</p> <p>-</p> <p>Samtidigt var der et ønske om beslutningsstøtte til børn.</p> <p>-</p> <p>Forskellige interview gav forskellige billede af hvad der var ønske om.</p>	<p>Praktiserende læger vil få mest ud af beslutningsstøtte til gravide og ammende og besværlige medicinske patienter.</p> <p>-</p> <p>Læger vil være sikre på at forskelligt medicin kan spille sammen.</p> <p>-</p> <p>Læger skal interessere sig for patienterne. Teknologien skal være ligevalgt, så længe det kan hjælpe patienter. Her skal læger mødes på halvvejen.</p> <p>-</p> <p>Mange parametre skal vurderes i systemet. Hvilken alder, hvornår kan man være gravid, mand vs. kvinde.</p> <p>-</p> <p>Et ønske om beslutningsstøtte til behandling af børn.</p> <p>-</p>

Continued on next page

Table C.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analyse og fortolkning
08:57-30:40 (fortsat)	<p>Gav også mere information til hvilke spørgsmål/parametre der skal opfylDES for gravide/ammende.</p> <p>-</p> <p>Hvor langt man er i graviditet/hvor lang tid har man ammet.</p> <p>-</p> <p>Systemet skal indenfor de næste to måneder kunne give svar på data der skal ind i systemet.</p> <p>-</p> <p>Alfa Omega er at man har tillid til det svar der kommer ud fra systemet.</p> <p>-</p> <p>Der er stramme krav til hvor datagrundlaget kommer fra, hvem har lavet det og hvilke organisation er det? DLI er privat, men en meget stærk videnskabelig organisation.</p> <p>-</p> <p>Kan man være sikre på de er uvildige?</p> <p>-</p> <p>Tilliden til systemet skal testes noget mere, og der skal kunne besvares hvorfor systemet svarer som det gør.</p> <p>-</p>	<p>Alfa Omega er tillid til systemet. Lægerne skal stole på det råd der kommer fra systemet</p> <p>-</p> <p>Stramme krav til datagrundlag. Organisationen skal være uvildig, selvom DLI er privat men en videnskabelig organisation.</p> <p>-</p> <p>Systemet skal kunne besvare hvorfor det giver rådet. Ligner meget samme tendens som set i tidligere interview. Transparens er vigtig, da læger kan sætte spørgsmålstegn ved systemet og lære.</p>

Continued on next page

Table C.2 – continued from previous page

Interviewsekvens	Meningskondensering	Analysé og fortolkning
30:42-38:59	<p>Trifork leverer 11 moduler, med forskellige fokus.</p> <ul style="list-style-type: none"> - - Gravide og ammende, Diagnoser, Maksimum, gfr-værdi, antibiotikavejledning, etc. - Nogle ting stammer fra FMK og det de har lært der. Bl.a. advarsler mod dobbeltordination. - Der findes andre datagrundlag der er bedre, men hvis man er ude i verden skal alle data omkonverteres. - Lægerne vil gerne have beslutningsstøtte men ikke hvis det kræver ekstra arbejde. - Der er blevet slået fast at lægens ansvar ikke fjernes. Systemet skal bare give råd og lægen vil altid have ansvaret stadigvæk for hvad der sker. - Lægerne er ikke så tit blevet spurgt omkring udviklinge af nye systemer. - De praktiserende lægers organisation er med i udviklingen denne gang. - Meget diskussion omkring hvordan de beviser at der er tillid i systemet, så istedet prøver de at opretholde de krav der er blevet sat til systemet og de brugerkrav der blev udformet. 	<p>11 forskellige moduler. Hver modul har forskellig fokus. fx. gravide og ammende, diagnoser, maksimum, antibiotikavejledning.</p> <ul style="list-style-type: none"> - Dobbeltordinationsadvarsler. Flere ting er lært fra FMK implementeringen - Omkonvertering er besværligt. Kan være en grund til at DLI også er valgt. Læger vil gerne have beslutningsstøtte. Dog ikke hvis det kræver ekstra arbejde. Det skal passer til deres arbejde, passer overens med tidligere interview. - Lægerne beholder alt ansvar. Systemet kan give råd men tager ikke ansvar for rådet. - De praktiserende lægers organisation er med i udviklingen. Læger bliver sjældent involveret i udviklingen, dette er dog ændret i udviklingen af beslutningsstøtte. - Lægerne skal have tillid i systemet. Hvordan finder de ud af om der er tillid til systemet/rådet? De prøver først og fremmest at opretholde de brugerkrav der er til systemet.

Contextual interview notes

D

This appendix contains the written notes from the contextual interview. Firstly, a list of notes from one of the interviewers, and secondly a list of notes from the other interviewer. Lastly the appendix contains the notes from the interpretation session.

First interviewer:

- Det hele bliver kontrolleret
- Han selv har mange bolde i luften ad gangen
- Tænker selv abstrakt
- De har en clinical suite der samler alle deres systemer under et single sign-on
- Kan oprette en ordination
- Kontekstafhængige lister til den enkelte læge
- Ved ikke hvad der finder lægemidler når man søger
- Han bruger sin viden fra hovedet, eg. tag vand ud af kroppen.
- Problemorienteret -> Hvor man kan vælge forskellige specialer -> Hvor man kan skrive en problemformulering.
- Objektive undersøgelser -> Vægt -> Blodtryk -> Vand -> osv.
- Problem at ikke alt kan logges
- De kan se grafer i systemet
- Når han er ved at ordinere vil han gerne vide at det påvirker forskellige ting. Han vil gerne have tal.
- Dataene kan være der med det samme.
- Han ser patienten -> Blodtryk -> Hævede ben -> Værdier -> Etc.
- Han har en fornemmelse (Agressive - patient)
- Beslutningsstøtte er vigtig - man skal bare få dataene
- Han kan godt lide det, hvis det giver ham noget information præsenteret
- Han bruger interaktioner, men i hovedet
- Der var en alarm som var god, i dag kommer der en der er knap så god.
- Den kvikke sygeplejerske hjælper også med beslutningsstøtte
- Beslutningsstøtte skal advare, men ikke konstant
- De sparer lægerne imellem hele tiden
- Det foregår ved konferencer, de sidder på kontorer ude i afdelingerne, her hører de hinanden ad, for eksempel, hvis der er noget der ikke virker
- Vil gerne have de vigtigste værdier når han åbner en patient
- Vil også gerne have information om, hvem der snart skal hjem, eg. Læsø-færgen.
- Det kan godt blive for meget rødt lys
- Det skal være klinisk meningsfuldt
- Det skal være smart
- Det skal ikke styre ham

Second interviewer:

- Data ligger samlet i program (Clinical Suite) - Det giver overblik
- Det er aldrig ligetil
- Vil gerne have beslutningstøtte i form af huskeliste
- Vil gerne have et overblik og vurdering af de kritiske patienter i systemet
- Vil ikke have en ny arbejdsgang
- Det skal være klinisk meningsfuldt
- Aldrig ekstra arbejdsgange/systemer men altid sorg for at gøre det en del af det nuværende

Interpretation session:

- Region Nordjylland vil kontrollere alt, gøre det simpelt, sætte det i kasser

- Systemet skal ikke kontrollere ham, men omvendt
 - Systemet "nedgør" hans faglighed
- Han tænker meget abstrakt og har det meste i hovedet
 - Han har en masse information om hvad hans erfaring er og hvad han tidligere har lært.
- Clinical suite til at samle alle programmerne
 - Medicinordinering Columna viser en liste over medikamenter men ikke i nogen specifik orden
 - Her kan den med fordel se på hvem lægen er (Bruger) eller hvem patienten er og vise en liste over mest anvendte medikamenter.
- Sætte patinter i problemformuleringer der kan vælges. fx. "Patienten har for høj væskebalance" og derved kommer systemet med de mest relevante forslag el. værdier der skal tages højde for.
 - Vælge mellem forskellige speciale (Hjertekirugisk, Lunge, Tarm, osv.)
 - * Så kommer forskellige problemformuleringer frem eller forslag
- Når der ordineres vil han gerne vide hvad det kan interagere med og hvad det ikke kan. Der vil der gerne haves værdier og grafer præsenteret.
 - Eksempel: "Du er ved at give et vanddrivende stof, har du husket at kigge på nyrefunktionsværdien?" - Præsentation af værdien så lægen kan se hvorvidt den er god nok.
- Lægen danner sig et helt overblik og bruger fornemmelse til at vurdere patienten og mængden af medicin der skal bruges.
 - Eksempel "Fornemmelse bruges til at vurdere hvorvidt patienten skal have 30, 40 eller 50 mg af et bestemt stof. Svært at sætte om til 0 og 1 taller (Programmering)"
- Fejl kan ske fordi man glemmer at tjekke andre værdier der skal tjekkes, her kan sygeplejerske ofte hjælpe med at huske. Det samme ville et beslutningsstøtte system
- Læger sparrer ofte imellem hinanden på konference og kontorer (Sker ofte), med at vurdere en patient eller ordinering af medicin.
- Beslutningsstøtte til at vurdere hvilke patienter der er mest kritiske og hvem der skal behandle først. Samt have styr på hvordan patienten kan komme hjem (Fx. toget går kun kl. 12, derfor skal de vurderes og være ude før)
- Samtidigt et overblik over patienten i stedet for at skulle se alle faner igennem for at finde de vigtige værdier.
- Det kan godt blive lidt for meget rødt lys
- Det skal være klinisk meningsfuldt
- Det skal ikke være en ekstra arbejdsgang.
 - Systemet skal tilpasses dem og ikke dem der skal tilpasse sig systemet
- Huskelisten, er nævnt tidligere, men skal vise lægen om kritiske værdier der skal tjekkes der skal tages højde for inden ordination.
- Det skal aldrig være et ekstra system men i stedet en del af hvordan de arbejder nu.
- Han vil gerne guides af beslutningstøtten gennem men ikke afgrænsset af systemet.
 - Hvis systemet kan designes så det virker som om det er en anden læge der sidder med ham og hjælper ham som de plejer at gøre.
 - Eksempel - Udfylde ordination af medicin og indtastning af forkert værdi så skal systemet evt gøre ham opmærksom men ikke begrænse ham (skriver 100 istedet for 10) - (Som hvis en anden læge sad og hjalp ham med at vurdere og ordinere)

Medical Information

E

This appendix contains the medical information used in the medical history as scenario for the experimental design. The medical information can be seen in [Figure E.1](#).

Kompleks ældre medicinsk patient						
Køn	Mand					
Alder eller fødselsdato	78 år					
Højde	184 cm					
Vægt	95 kg					
BMI	28					
Legemsoverflade						
Diagnose	Type 2 diabetes, Hypertension, Tidl. AMI, tidl. Apopleksi, KOL					
Diagnosekode	T90, K86, K75, K90, R95					
Evt. Cave	Ingen					
Laboratorieværdier						
	Dato	Dato	Dato	Dato		
	27.12.18	04.05.19				
Kost	God					
Rygning	Daglig					
Alkohol (genstande)	14/uge					
Motion (timer)	2/uge					
Systolisk blodtryk	157	195				
Diastolisk blodtryk	70	95				
EKG	Taget					
Blodsukker	6,2	10,2				
Hgb	9,1					
Natrium	139					
Kalium	4,2					
GFR	71					
Creatinin	91					
ALAT	15					
Total kolesterol	4,9					
Triglycerider	0,87					
LDL kolesterol	2,9					
HDL kolesterol	1,7					
Hba1c	37					
U-alb	8					
U-crea	13,3					
U-alb/crea	5					
Andre relevante lab tal						
	Dato	Dato	Dato	Dato		
	20-08-2018					
FEV1	1,98					
FVC	2,96					
FEV1/FVC	0,66					
FVC%	69					
FEV1%	62					

Figure E.1: Medical information used to create the scenario for the experimental study.

Review session

F

This appendix contains a detailed description of the review session made in order to finalise the pairwise comparison study, the statements, the character descriptions and the overall scenario and context.

F.1 Session design

The main goal of this review is to get feedback on the medical scenario and the descriptions of the adviser characters and to make sure that they fit into a medical context.

F.1.1 Participant

The review is conducted as a small session with a medicine student from Aalborg University. The participant is a fourth semester bachelor student of medicine at Aalborg University, as well as a student worker at a local group medical practice with 1 years experience.

F.1.2 Setup and environment

The following subsections describe the materials of the pairwise experiment as well as the experimental setup.

Materials

- Lenovo laptop containing the experimental study to provide context
- Application with pairwise comparison experiment, see [Appendix I](#)
- Descriptions of characters (on paper)
- Description of a scenario (on paper)

Session setup

The session is carried out at the location of the subject, this means that the session is carried out at the apartment of the participant. The computer is set up with the program running, to provide the participant with the context of the pairwise experiment.

F.1.3 Session procedure

The procedure for the review session is as follows:

1. Introduction to the study and context by showing the pairwise comparison study, see [Appendix I](#).
 - a) The participant is asked whether the task of pairwise comparison is understandable.
 - b) The participant is asked for comments about the experimental setup.
2. The participant is introduced to a list of the statements [subsection 4.4.4](#).
 - a) The participant is asked whether the statements make sense.
 - b) The participant is asked for comments about the statements.
3. The participant is introduced to the scenario, see [subsection 4.4.6](#).
 - a) The participant is asked whether the medical scenario makes sense.
 - b) The participant is asked for comments about the scenario.
4. The participant is introduced to the adviser character descriptions, see [Appendix H](#) and [Appendix G](#).
 - a) The participant is asked whether the descriptions are accurate and fits the context.
 - b) The participant is asked for comments about the character descriptions.

During the session notes are written down and supplementary questions are asked in order to interpret the key points made by the participant.

F.2 Results and changes

During the session the most important key points and changes are as follows:

There were changes to sentences to make them better fit the medical context. An example is "the patient has hypertension, type II diabetes, AMI..." was changed to "the patient is under treatment for hypertension type II diabetes, and has previously had AMI and...". Furthermore, the addition of laboratory values from the patient of the scenario are include based on the feed-back from the session. Lastly, it was chosen to rename the "scenario" to be called a "medical case history" (sygehistorie), since this is the term used by clinicians and consultants when they work with patients.

The additions made to the characters include: A more detailed description of "PrescriptAid", for instance where the data comes from, pro.medicin, who has made the system, which values the advice is based on an whether it is automatically updated. And a more detailed description of the consultant, the name is changed from being a generic name, Jens Jensen, to another Danish name, Asger Knudsen, furthermore, the consultant is to a larger extent described with medical terms, and his medical education and years of experience was mentioned instead of more personal features such as wife, children, where he lives, and that he has an interest in technology.

Consultant

G

This appendix contains the description of the revised consultant as used in the experimental study. The description is printed to a full-sized A4 sheet of paper.

Overlæge Asger Knudsen



Profil

Alder: 62 år

Organisation: Region Nordjylland

Køn: Mand

Ansættelsested: Aalborg Universitetshospital

Status: Gift, fire børnebørn

Job: Ledende overlæge på kardiologisk afdeling

Lokation: Aalborg

Stammer fra Nordjylland og bor i Hjørring med sin kone og hund. Asger har gennem hele sin karriere specialiseret sig indenfor kardiologi. Han er uddannet fra Aarhus Universitet i 1983, og har tidligere været afdelingslæge på kardiologisk afdeling hos Skejby Sygehus men er nu ledende overlæge på kardiologisk afdeling hos Aalborg Universitetshospital.

Råd

Gennem din diagnosticering fastslår du at ordinere Rasilez til patienten men vælger at få en vurdering af din beslutning fra din kollega Asger Knudsen. Din kollega nævner at patienten er diagnosticeret med diabetes og derfor ikke burde få Rasilez men istedet en forhøjet dosis af den nuværende blodtryksmedicin, Cozaar.

Figure G.1: shows the description of the consultant, *Asger Knudsen*, as used in the experimental study.

PrescriptAid

H

This appendix contains the description of the revised decision aid as used in the experimental study. The description is printed to a full-sized A4 sheet of paper.

PrescriptAid

PrescriptAid - Medicinordning

Præparatets navn: Generisk navn:

Rasilez Indtast søgeord

Præparatets navn	Generisk navn	Form
Rasilez	Aliskiren	Tablet
Cozaar	Losartan	Tablet
Ancozan	Losartan	Tablet
Klomentan	Losartan	Tablet

Undgå Rasilez, forøg eventuelt dosis af Cozaar i stedet

Støtte til beslutning:
Diagnos: Type 2 diabetes
Interaktioner med medicamenter: Metformin "Actavis", Cozaar
Maksdosis: Cozaar nuværende: 50 mg, 1tbl. dgl.
Cozaar maksdosis: 100 mg, 1tbl. dgl.

Aliskiren (Rasilez) bør ikke ordnes til patienter med diabetes
Patienten behandles i forvejen med Cozaar 50 mg, 1 tablet dagligt
Ordner eventuelt forøget mængde af Cozaar, op til 100 mg, 1tbl. dgl.

Viser 1 af 1 samt 3 af 3 lignende præparer

ANULLÉR ORDINER

Profil

Version: 4.0.3

Datagrundlag: pro.medicin

Skabt af: PharmaIT A/S

Definition: Beslutningsstøttesystem

Organisation: Region Nordjylland

Funktion: Hjælpe med ordning af medicin

PrescriptAid er et beslutningsstøttesystem udviklet af PharmaIT A/S i Aalborg. Systemet anvendes primært af læger på sygehuse samt praktiserende læger. Systemet er udviklet til at hjælpe lægen med ordination af medicin, med det formål at nedsætte fejlordinationer. Rådende er baseret på nyeste viden fra pro.medicin om maksdoser, interaktioner mellem medicin og allergier. PrescriptAid opdateres løbende med information om patienten, for eksempel laboratoriesvar.

Råd

Gennem din diagnosticering fastslår du at ordinere Rasilez til patienten. PrescriptAid gør opmærksom på, at patienten er diagnosticeret med diabetes og spørger, hvorvidt du stadig vil ordinere Rasilez, da patienter med diabetes ikke burde få det. Der foreslås i stedet at forhøje dosen af den nuværende blodtryksmedicin, Cozaar.

Figure H.1: shows the description of the decision aid, *PrescriptAid*, as used in the experimental study. A full-size image of the decision aid can be seen in Section 4.4.6

Pairwise software description

I

This appendix contains a description of the processes undergone during development of the experimental software for the pairwise comparison experiment. The following sections describe the programming and development processes, from preparation and planning, to the objects and functions, and lastly the actual user interface and functionality of the experimental software.

I.1 Preparation and planning

Since the purpose of the software is to gather experimental data from a pairwise comparison experiment, the software was developed with three states; an introduction state, the pairwise comparison state, and an end state. A flow diagram illustrating the three states can be seen in Figure I.1.

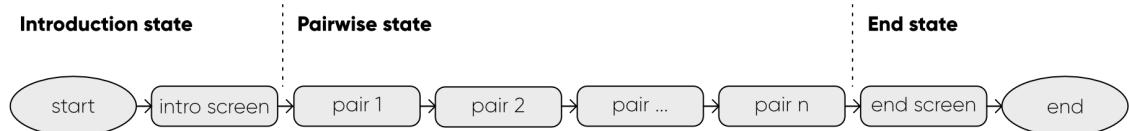


Figure I.1: Shows a flow diagram for the developed experimental software.

The initial plan for the experimental study was a combined study consisting of both a pairwise comparison and a scaling experiment. Hence, the original software as developed for this experiment consisted of two additional states; an attribute scaling state consisting of visual analogue scales for all statements for each character and a preference scaling state consisting of visual analogue scales for a preference of advice from each character. The two scaling experiment states was developed such that they could be initialised previously to or after the pairwise state. These are not explained since they were not used in the final experimental design due to time limitations and difficulty for participants to rate the statements on visual analogue scales during the pilot.

I.2 The code

In order to simplify the development process it was chosen to use object-oriented programming. This also simplifies adding new buttons and pages/screens without having to start from the bottom each time. This subsection describes the objects and functions developed for the experimental software. The programming language chosen for programming the experimental software is Processing 3.5.3 for Windows 64-bit ([Fry & Reas, 2019](#)). The program can also be seen in the external appendix as a zip-file, see [section L.1](#).

I.2.1 Global variables

Firstly the software contains a list of global variables. These were mainly used for fonts, initialising objects, the three states, tables for importing and outputting data (characters, statements and results) from and to .csv files, Strings for on-screen text, and variables for the output file. The variables are seen in the following code:

```
//Fonts  
PFont GilroyLightFont;  
PFont GilroyBoldFont;  
  
//The objects  
Button startButton;
```

```

Button endButton;
Button2 pairwiseButton1;
Button2 pairwiseButton2;
PageBackground introPageBG;
TransitionScreen beginScreen;
TransitionScreen endScreen;
PairScreen[] PairChoice;

//The three states
boolean pairState = false;
boolean pressedBefore = false;
boolean introPageState;
boolean endState = false;
boolean pairClicked = false;

//The tables containing the characters, statements and a table for saving the results
Table charTablePair;
Table statTablePair;
Table PairResults;

//Strings for characters and statements for presenting them on screen and for
//saving them to the output file
String[] ChaPair;
String[] StatPair;
String[] CharStat;
String[] CharStatSave;
String[] CharStatOppSave;
String[] PairPres1Save;
String[] PairPres2Save;
String[] PairPres1;
String[] PairPres2;

//Variables for calculating amount of presentations and for saving values to the
//output table
int statAmount;
int answersGiven;
int[] answ;
int[] answPair;
int[] answOppPair;
int chaAmountPair;
int presAmountPair;

//Variables that change for each subject. fp is changed before each subject
//starts the experiment. Random number is used to avoid deleting data in
//the case that a change of 'fp' is forgotten.
int fp = 1;
int randomNum=int(random(999999));
String charAttr;
int index = 0;

//Text for screens
String introductionText = "Velkommen og tak fordi du har lyst til at deltage";
String initialDescription = "Når du har fået en kort, mundtlig introduktion til
forsøget, og du har \n gennemlest og underskrevet samtykkeerklæringen,
er vi klar til at begynde. \n \n \n Tryk på 'start' for at begynde.";
String endText = "Forsøget er nu ovre, tak for din deltagelse";
String endDescription = "Henvend dig venligst til facilitatoren";

```

I.2.2 Objects, classes and functions

The following subsubsections contains a small explanation of each class and function.

The background class

The first class created is the background. This background is drawn in order to refresh the whole page and be able to draw new pages on the screen.

```

class PageBackground {
    color c;
    int xpos;

```

```

int ypos;
int w;
int h;

PageBackground (color tempC, int tempXpos, int tempYpos, int tempW, int tempH) {
    c =tempC;
    xpos = tempXpos;
    ypos = tempYpos;
    w = tempW;
    h = tempH;
}

void displayPageBackground() {
rectMode(CENTER);
fill(c);
noStroke();
rect(xpos, ypos, w, h);
}
}

```

The shadow effect function

For the background and in order to highlight buttons on the screen, a shadow effect is created:

```

void glowingRect(float x, float y, float w, float h, color rectColor, color glowColor) {
    noFill();
    for (int i=0; i<5; i++) {
        strokeWeight((2+i)*(2+i)/5);
        stroke(glowColor, 10);
        rectMode(CENTER);
        rect(x, y, w, h);
    }
    noStroke();
    fill(rectColor);
    rect(x, y, w, h);
}

```

The transition screen class (begin/end)

In order to create an introduction and ending screen a class named transitionScreen is created.

```

class TransitionScreen {
    String heading;
    int xposHead;
    int yposHead;
    String description;
    int xposDesc;
    int yposDesc;

    TransitionScreen (String tempHeading, int tempXposHead, int tempYposHead,
    String tempDescription, int tempXposDesc, int tempYposDesc) {
        heading = tempHeading;
        xposHead = tempXposHead;
        yposHead = tempYposHead;
        description = tempDescription;
        xposDesc = tempXposDesc;
        yposDesc = tempYposDesc;
    }

    void displayTransitionScreen() {
        GilroyLightFont = loadFont("Gilroy-Light-48.vlw");
        GilroyBoldFont = loadFont("Gilroy-ExtraBold-48.vlw");
        noStroke();
        fill(102);

        textAlign(CENTER);
        text(heading, xposHead, yposHead);
    }
}

```

```

    textFont(GilroyLightFont, 24);
    textAlign(CENTER);
    text(description, xposDesc, yposDesc);
}
}

```

Two Button classes

In order to create buttons for the introduction screen as well as buttons for the pairwise comparison screen, two different button classes are created. This could have been avoided by not hard-coding the text-sizes of the buttons into the actual button classes. This could be refactored if the program is to be used going forwards. Hence, the only thing differentiating the two button classes are the text-sizes.

```

class Button {
    String butText;
    int xpos;
    int ypos;
    color c;
    color cText;
    color shadowColor;
    boolean dropShadow;
    int w;
    int h;

    Button (color tempC, color tempCtext, int tempXpos, int tempYpos, String tempButText,
            color tempShadowC, boolean tempDropShadow, int tempW, int tempH) {
        c = tempC;
        cText = tempCtext;
        xpos = tempXpos;
        ypos = tempYpos;
        butText = tempButText;
        shadowColor = tempShadowC;
        dropShadow = tempDropShadow;
        w = tempW;
        h = tempH;
    }

    void displayButton() {
        noFill();

        if (dropShadow == true) {
            for (int i=0; i<5; i++) {
                strokeWeight((2+i)*(2+i)/5);
                stroke(shadowColor, 20);
                rectMode(CENTER);
                rect(xpos, ypos, w, h, 7);
            }
            noStroke();
        }
        rectMode(CENTER);
        fill(c);
        rect(xpos, ypos, w, h, 7);

        GilroyLightFont = loadFont("Gilroy-Light-48.vlw");
        fill(cText);
        textFont(GilroyLightFont, 36);
        textAlign(CENTER,CENTER);
        text(butText, xpos, ypos);
    }
}

class Button2 {
    String butText;
    int xpos;
    int ypos;
    color c;
    color cText;
    color shadowColor;
    boolean dropShadow;
}

```

```

int w;
int h;

Button2 (color tempC, color tempCtext, int tempXpos, int tempYpos,
String tempButText, color tempShadowC, boolean tempDropShadow, int tempW, int tempH) {
    c = tempC;
    cText = tempCtext;
    xpos = tempXpos;
    ypos = tempYpos;
    butText = tempButText;
    shadowColor = tempShadowC;
    dropShadow = tempDropShadow;
    w = tempW;
    h = tempH;
}

void displayButton() {
    noFill();

    if (dropShadow == true) {
        for (int i=0; i<5; i++) {
            strokeWeight((2+i)*(2+i)/5);
            stroke(shadowColor, 20);
            rectMode(CENTER);
            rect(xpos, ypos, w, h, 7);
        }
        noStroke();
    }
    rectMode(CENTER);
    fill(c);
    rect(xpos, ypos, w, h, 7);
}

GilroyLightFont = loadFont("Gilroy-Light-48.vlw");
fill(cText);
textFont(GilroyLightFont, 24);
textAlign(CENTER,CENTER);
text(butText, xpos, ypos);
}
}

```

Pairwise comparison screen class

To display the pairwise comparison screen for the participants a class called pairScreen is created.

```

class PairScreen {
    String Description;
    int xposDesc;
    int yposDesc;
    String charStat1;
    String charStat2;

    PairScreen (String tempDesc, int tempXposDesc, int tempYposDesc, String tempCharStat1,
    String tempStatement) {
        Description = tempDesc;
        xposDesc = tempXposDesc;
        yposDesc = tempYposDesc;
        charStat1 = tempCharStat1;
        charStat2 = tempStatement;
    }

    void displayPairScreen() {
        GilroyLightFont = loadFont("Gilroy-Light-48.vlw");
        GilroyBoldFont = loadFont("Gilroy-ExtraBold-48.vlw");
        fill(255);
        rect(width/2, height/2, width/6*5, height);
        noStroke();
        fill(102);
        textFont(GilroyLightFont, 48);
        textAlign(CENTER);
        text(Description, xposDesc, yposDesc);
    }
}

```

```

pairwiseButton1 = new Button2(#FFFFFF, 102, width/2, height/6*3, charStat1,
#000000, true, 1200, 150);
pairwiseButton2 = new Button2(#FFFFFF, 102, width/2, height/6*4, charStat2,
#000000, true, 1200, 150);
}
}

```

Pairwise experiment function

In order to load the characters and statements to be pairwise compared, a function is used. This function loads the characters and statements from external files, combines them, creates an output table, makes sure pairs are only shown once, randomises the statements, and creates a screen object for each comparison of pairs.

```

void pairwise () {
    charTablePair = loadTable("Characters.ods", "header");
    statTablePair = loadTable("Statements.ods", "header");

    PairResults = new Table();

    chaAmount = charTablePair.getRowCount(); //Værdier hentes fra odt filer
    statAmount = statTablePair.getRowCount(); // Værdier hentes fra odt filer
    charStatAmount = (statAmount*chaAmount);
    presAmountPair = (((statAmount*chaAmount)-1)*(statAmount*chaAmount))/2;

    ChaPair = new String[chaAmount];
    StatPair = new String [statAmount];
    CharStat = new String [charStatAmount];
    PairPres1 = new String [presAmountPair];
    PairPres2 = new String [presAmountPair];
    CharStatSave = new String [charStatAmount];
    CharStatOppSave = new String [charStatAmount];
    PairPres1Save = new String [presAmountPair];
    PairPres2Save = new String [presAmountPair];
    PairChoice = new PairScreen[presAmountPair];

    answPair = new int[presAmountPair];
    answOppPair = new int[presAmountPair];
    println(presAmountPair);

    PairResults.addColumn();

    //Værdier hentes fra tables:
    int i = 0;
    int j = 0;
    for (TableRow chars : charTablePair.rows()){
        ChaPair[j] = chars.getString("Karakter");
        int k=0;
        for (TableRow stats : statTablePair.rows()) {
            StatPair[k] = stats.getString("Statement");
            CharStatSave[i] = ChaPair[j] + " " + StatPair[k];
            CharStat[i] = ChaPair[j] + "...\\n\\n..." + StatPair[k];
            CharStatOppSave[i] = ChaPair[j] + " " + StatPair[k];

            PairResults.addColumn(ChaPair[j] + " " + StatPair[k]);
            TableRow newRow = PairResults.addRow();
            newRow.setString(0, ChaPair[j] + " " + StatPair[k]);
            i++;
        }
    }

    //karakter og statement opstilles parvis så de kun sammenlignes en gang med hinanden samt
    at de ikke sammenlignes med sig selv , i to arrays.
    i=0;
    for (int l = 0; l < CharStat.length; l++) {
        for (int m = l + 1; m < CharStat.length; m++) {
            if (CharStat[l] != CharStat[m]) {
                //do stuff
                int randomizer=int(random(2));

```

```

        println("Random value: " + randomizer);
        if(randomizer == 0) {
            PairPres1Save[i] = CharStatSave[1];
            PairPres2Save[i] = CharStatSave[m];
            PairPres1[i] = CharStat[1];
            PairPres2[i] = CharStat[m];
        } else if (randomizer == 1) {
            PairPres2Save[i] = CharStatSave[1];
            PairPres1Save[i] = CharStatSave[m];
            PairPres2[i] = CharStat[1];
            PairPres1[i] = CharStat[m];
        }
        i++;
    }
}

//De to arrays randomiseres
println(PairPres1.length);
//println(PairPres1[9]);
for (int l = PairPres1.length-1; l > 0; l--) {

    // Pick a random index from 0 to l
    int m = int(random(l+1));

    // Swap arr[l] with the element at random index
    String temp2 = PairPres1[l];
    String temp3 = PairPres2[l];
    PairPres1[l] = PairPres1[m];
    PairPres2[l] = PairPres2[m];
    PairPres1[m] = temp2;
    PairPres2[m] = temp3;

    String temp2Save = PairPres1Save[l];
    String temp3Save = PairPres2Save[l];
    PairPres1Save[l] = PairPres1Save[m];
    PairPres2Save[l] = PairPres2Save[m];
    PairPres1Save[m] = temp2Save;
    PairPres2Save[m] = temp3Save;
}

//Der laves skærbilleder for hver parvis sammenligning:
for (int l= 0; l < PairPres1.length; l++){
    PairChoice[l] = new PairScreen("Vælg den af de to muligheder, der er mest rigtig",
    width/2, height/6*2, PairPres1[l], PairPres2[l]);
}
}

```

Save function

In order to save the collected data to an output file, the following function is created. The output file is saved as a csv.

```

// funktion til at gemme data:
void SaveDataPair() {
    saveTable(PairResults, "Data/FP"+fp+" - "+randomNum+".csv", "csv");
    println("Data is saved");
}

```

The setup function

In processing `void setup` is used for code, that is run only once. Hence, here the button objects, background object, and introduction and end screen objects are created. Furthermore, the introduction screen is displayed, the *pairwise* function is run and the start button is displayed.

```

void setup() {
    fullScreen();
    background(#E5E5E5);
}

```

```

endButton = new Button(#2E7D32, #FFFFFF, width/2, height/6*5, "afslut",
#000000, true, 130, 90); //Lav slut-knap
startButton = new Button(#2E7D32, #FFFFFF, width/2, height/6*5, "start",
#000000, true, 130, 90); // Lav start-knap
betweenExpButton = new Button(#2E7D32, #FFFFFF, width/2, height/6*5, "næste",
#000000, true, 130, 90); //Lav knap til mellem eksperimenter

introPageBG = new PageBackground(255, width/2, height/2, width/6*5,
height); // lav baggrund
beginScreen = new TransitionScreen(introductionText, width/2,
height/6, initialDescription, width/2, height/6*2); // lav startskæm
betweenExpScreen = new TransitionScreen(betweenExpText, width/2,
height/6*2, betweenExpDesc, width/2, height/6*3); //lav skæm til mellem eksperimenter
endScreen = new TransitionScreen(endText, width/2, height/6*2,
endDescription, width/2, height/6*3); // lav slut-skæm

introPageState = true; //sæt intro stadie

glowingRect(width/2, height/2, width/6*5, height, #FFFFFF, #000000); //lav drop-shadow til baggrund

introPageBG.displayPageBackground(); //tegn baggrund
beginScreen.displayTransitionScreen(); // tegn startskæm

pairwise(); // Load statements og karakterer, samt randomiser så de er klar
til at kunne præsenteres pairwise.
startButton.displayButton(); // Vis startknap
}

```

The draw function

In processing `void draw` is used for code that is run repeatedly. This function runs the code that displays the pairwise screen matching the pairwise comparison that has to be answered. Each time a new answer is made the value `answersGiven` increases by one and the next pairwise comparison is shown until the amount of answers given is equal to the length of the array of presentation. Hence, this section contains the following code:

```

void draw() {

    if (pairState == true && answersGiven < PairPres1.length) {
        introPageBG.displayPageBackground();
        PairChoice[answersGiven].displayPairScreen();
        pairwiseButton1.displayButton();
        pairwiseButton2.displayButton();
    }
}

```

Mouse actions

Since the software is run on a Windows PC with touch the following code is necessary in order to be able to use the touch-screen capabilities. Firstly, if the start-button is pressed the pairwise state begins. Secondly, it is now registered whether the answer given on the screen is the first statement or the second statement in the pairwise comparison. The choice is registered to the result table. This is repeated until the amount of answers given are one smaller than the amount of presentations. When this is the case, the mouse click saves the last choice, creates the output file, and shows the end screen.

```

void mouseReleased() {
    if ((introPageState == true) && (mouseX >= startButton.xpos - 150/2) &&
(mouseX <= startButton.xpos + 150/2) && (mouseY >= startButton.ypos - 100/2) &&
(mouseY <= startButton.ypos + 100/2)) { //IntroductionScreen
        introPageState = false;
        pairState = true;
    }

    else if (pairState == true) {
        if (mouseY>=pairwiseButton1.ypos - (pairwiseButton1.h/2) &&

```

I.3 User interface and functionality

In the introduction state the experimental participant is shown on Figure I.2. When the participant presses the 'start' button they are presented with the pairwise comparisons that are imported from .csv files containing the characters and statements. In the current experiment the 2 characters and 10 statements combined gives a total of 190 pairwise comparisons. Three examples of these pairwise comparisons are seen on Figures I.3, I.4, and I.5. When the par-

ticipants have made the total amount of pairwise comparisons they are presented with the end screen as seen in I.6.



Figure I.2: Shows the introduction screen for the pairwise comparison experiment.



Figure I.3: Shows an example of a pairwise comparison from the actual pairwise comparison experiment.



Figure I.4: Shows an example of a pairwise comparison from the actual pairwise comparison experiment.



Figure I.5: Shows an example of a pairwise comparison from the actual pairwise comparison experiment.



Figure I.6: Shows the end screen for the pairwise comparison experiment.

Pilot experiment

This appendix contains a description of the pilot for the experimental study carried out as part of this thesis. The appendix contains the sections for the pilot experiment, which differ from the final experimental design. Hence, sections and parts omitted from this appendix are to be found in the actual experimental design in [chapter 5](#). The sections that are found in this section are the *study description*, *method considerations*, as well as the experimental design sections of method description, setup and environment, the experimental procedure.

J.0.1 Experimental subjects

The pairwise pilot experiment is conducted with 5 subjects, with a mean age of 23.8 and a range of 21–25 years old, 1 medical students, 21 years old, and 4 university students from different degrees other than medicine mean age 24.5 and a range of 24–25 years old.

J.0.2 Setup and environment

The following subsections describe the materials of the pairwise experiment as well as the experimental setup.

Materials

- ASUS Zenbook UX 360U
- Application with pairwise comparison experiment, see [Appendix I](#)
- Descriptions of characters (on paper)
- Description of a scenario (on paper)
- Declaration of consent form (on paper)

Experimental setup

The experiment is carried out at the location of the subjects, this means that the experiment is carried out in Aalborg University group rooms with medical students and other students.

The computer is set up with the program already running in full screen. The subjects are able to perform the pairwise comparison study on the touch screen of the computer directly in the program.

The data from the experiment is saved locally onto the computer throughout the experiment as described in [Appendix I](#).

Stimuli

The stimuli presented to the participants are the 20 statements, 10 for each character, as well as the descriptions of the two adviser characters as described in [chapter 4 subsection 4.4.4](#) and [section 4.4.6](#).

J.0.3 Experimental procedure

The experimental procedure for the pairwise comparison experiment and exit interview is as follows:

1. Introduction to the study according to the script, see [Appendix K](#).
 - a) The participant is introduced to the scenario, see [subsection 4.4.6](#).
 - b) The participant is introduced to the adviser character descriptions, see [Appendix G](#) and [Appendix H](#).
2. The declaration of consent is signed on paper.

3. The pairwise comparison experiment begins.
 - a) An on-screen introduction screen is shown.
 - b) The pairwise comparisons are performed.
 - i. Two statements are compared at a time, the character descriptions are always present, and it is not allowed to return to previous comparison.
 - c) After 7 minutes a small break is held in which the participant is offered cake/cookies/fruit and coffee/tea/water.
 - d) The remaining pairwise comparisons are performed.
 - e) When the last comparison is made, the data is saved.
 - f) The participant is shown an end screen.
4. The exit interview begins.
 - a) One question is asked at a time according to the question list.
 - b) The answers are noted down in a google sheets spread sheet.
5. Debriefing is performed according to the script, [Appendix K](#).

J.1 Results

This section describes the experimental results of the pilot experiment, and more importantly the practical changes made based on the pilots of the experimental study.

J.1.1 Experimental results

The pilot experiment results in the following cumulative preference matrix based on the individual data sets pooled together, see [Table J.1](#). The rows and column Numbers are represented by the following character/statement combinations:

1. Overlæge Asger Knudsen giver råd baseret på videnskabelig evidens.
2. Overlæge Asger Knudsen giver råd baseret på faktorer der ikke er videnskabeligt testbare.
3. Overlæge Asger Knudsen giver råd baseret på et samlet billede af patienten
4. Overlæge Asger Knudsen er ansvarlig for sine råd
5. Overlæge Asger Knudsen kan forklare sit ræsonnement så det er forståeligt
6. Overlæge Asger Knudsen kan fremme ekspertise hos andre
7. Overlæge Asger Knudsen kan fremme egen ekspertise
8. Overlæge Asger Knudsen har et økonomisk incitament for sine råd
9. Overlæge Asger Knudsen tager højde for ændringer i patientens behov
10. Overlæge Asger Knudsen holder informationer fortrolige
11. PrescriptAid giver råd baseret på videnskabelig evidens.
12. PrescriptAid giver råd baseret på faktorer der ikke er videnskabeligt testbare.
13. PrescriptAid giver råd baseret på et samlet billede af patienten
14. PrescriptAid er ansvarlig for sine råd
15. PrescriptAid kan forklare sit ræsonnement så det er forståeligt
16. PrescriptAid kan fremme ekspertise hos andre
17. PrescriptAid kan fremme egen ekspertise
18. PrescriptAid har et økonomisk incitament for sine råd
19. PrescriptAid tager højde for ændringer i patientens behov
20. PrescriptAid holder informationer fortrolige

Table J.1: Shows the cumulative preference matrix for the pilot experiment consisting of the pooled data from 5 data sets.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-	3	2	3	1	2	3	3	1	4	0	5	5	4	1	0	4	3	4	3
2	2	-	0	3	0	1	0	4	0	1	0	5	3	2	0	1	3	4	2	3
3	3	5	-	4	4	4	3	4	4	5	2	5	3	3	2	4	4	5	5	4
4	2	2	1	-	2	1	2	5	0	2	2	4	2	4	2	1	4	4	2	2
5	4	5	1	3	-	2	4	4	1	4	1	5	5	4	2	2	5	4	4	3
6	3	4	1	4	3	-	3	5	2	3	2	5	4	4	2	2	5	5	5	3
7	2	5	2	3	1	2	-	4	2	3	1	4	4	3	1	2	4	4	3	4
8	2	1	1	0	1	0	1	-	0	1	0	3	1	2	0	1	2	3	2	1
9	4	5	1	5	4	3	3	5	-	3	2	5	4	5	2	3	5	5	4	4
10	1	4	0	3	1	2	2	4	2	-	1	5	3	4	2	2	4	3	2	3
11	5	5	3	3	4	3	4	5	3	4	-	5	4	4	5	4	4	5	4	4
12	0	0	0	1	0	0	1	2	0	0	0	-	1	2	0	2	0	1	1	0
13	0	2	2	3	0	1	1	4	1	2	1	4	-	5	1	1	4	3	1	1
14	1	3	2	1	1	1	2	3	0	1	1	3	0	-	1	2	2	4	2	1
15	4	5	3	3	3	3	4	5	3	3	0	5	4	4	-	3	4	4	5	5
16	5	4	1	4	3	3	3	4	2	3	1	3	4	3	2	-	4	4	5	5
17	1	2	1	1	0	0	1	3	0	1	1	5	1	3	1	1	-	2	3	2
18	2	1	0	1	1	0	1	2	0	2	0	4	2	1	1	1	3	-	0	0
19	1	3	0	3	1	0	2	3	1	3	1	4	4	3	0	0	2	5	-	3
20	2	2	1	3	2	2	1	4	1	2	1	5	4	4	0	0	3	5	2	-

Further data analysis was omitted for the pilot experiment.

J.1.2 Length of experiment

During the pilot test of the experiment the timestamps for the experiment were as follows: Approximately 6–7 minutes for introduction, approximately 21–25 minutes for pairwise comparison, and approximately 2–5 minutes of exit-interview. The total time for the pilot test ranged between 29 minutes and 20 seconds and 35 minutes and 22 seconds.

J.1.3 Exit interview results

During the pilot experiment it was found that one of the statements was difficult for the participants to understand. The statement was '...kan fremme egen ekspertise' (can enhance own expertise, especially in connection with the adviser character PrescriptAid). Hence, it is chosen to substitute this statement with the other statement regarding the same focus area '...kan omstille sig til ny viden' (can adapt to new information). When asked, the pilot experiment subjects found this wording more fitting to the context and easier to understand in connection with the PrescriptAid character adviser.

All subjects found it difficult to concentrate along the experiment, especially towards the end of the experiment after the break. Hence, it is chosen to change the break to occur every 10 minutes instead of only after 7 minutes. Furthermore, the participants sometimes found it difficult to choose between the options and therefore suggested that it is made clear that some choices can be difficult although it is necessary that they then choose the one they think is most correct. Hence, a part is added to the script stating that some pairwise comparisons might seem difficult.

One subject mentioned that there was a lot of reading to get into the right context and that this could be perhaps done orally instead. It was however chosen to keep the paper instructions to let participants get into the context by themselves, as well as make it possible for participants to use the characters for reference throughout the experiment.

Script

K

This appendix contains the script used in the experimental study.

Hej og velkommen til. Mit navn er Morten/Mikkel, og jeg vil facilitere forsøget i dag.

Vores projekt handler om beslutningsstøtte til læger, og derfor vil vi gerne have at netop du deltager. Som sagt skal du deltage i vores forsøg i dag, og her skal du tage stilling til nogle udsagn to ad gangen og vælge det, der er mest rigtigt.

Forsøget kommer til at foregå på denne computer, hvor du vil blive præsenteret for to udsagn af gangen, og hvor du så skal vælge det, der er mest rigtigt. Gennem forsøget skal du anvende touch på skærmen til at tilkendegive dit valg.

På et tidspunkt i forsøget stopper jeg dig, og så tager vi en lille pause, hvor du kan få noget at drikke eller tage et stykke kage. Du må selvfølgelig også gerne gemme det til efter forsøget.

Efter pausen fortsætter vi forsøget som før og når du er færdig vil skærmen vise dig det og du bedes henvende dig til mig igen. Tilsidst har jeg nogle korte spørgsmål til dig.

Hvis du er med på det, må du meget gerne underskrive denne samtykkeerklæring.

Jeg vil starte med at introducere dig for en sygehistorie, som du skal leve dig ind i og dernæst, vil jeg vise dig beskrivelser af to rådgivere, som du bedes gennemlæse.

“FP vises sygehistorie og rådgivere og gennemlæser dernæst sygehistorie og spørgsmål for sig selv.”

Hvad er dine spørgsmål?

“Besvar spørgsmål fra forsøgspersonen.”

Du skal være opmærksom på, at det nogle gange kan være besværligt at vælge mellem de to muligheder, men du skal i disse situationer stadig vælge den, du synes er mest rigtig, selvom at det kan være svært. Det er en del af forsøget.

“Forsøgspersonen udfører forsøget i 10 minutter”

Nu holder vi lige en lille pause, hvor du er velkommen til at tage lidt kaffe/te/vand og kage.

“Forsøgspersonen udfører anden halvdel af forsøget”

Electronic appendix overview L

This appendix shows a table of content for the attached electronic appendix. Each file will be found in the attached appendix labelled "ElectronicAppendix" which is a zipped file. Inside, folders named after each of the following titles are found.

L.1 PairwiseComparison

This electronic appendix contains a Processing file with the software used for the pairwise comparison study data collection.

L.2 RawData

This electronic appendix contains the raw data from the participants of the pairwise comparison study.

L.3 ExitInterview

This electronic appendix contains the exit interview notes from the participants of the pairwise comparison study.

L.4 Rscript

This electronic contains the R script used to carry out the data analysis of the data collected in the pairwise comparison study.