Evaluation of medication-related clinical decision support using user interface design principles
A qualitative study of Sfinx and NjuRen

Author: Emma Molin
Affirmation

I hereby affirm that this Master thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text. This work has not been submitted for any other degree or professional qualification except as specified; nor has it been published.

Stockholm, 2016-06-10

Emma Molin
Evaluation of medication-related clinical decision support using user interface design principles

A qualitative study of Sfinx and NjuRen

Abstract

Background: The process of prescribing medicine to patients has been recognized as an international issue due to the process being complex and error prone, causing preventable harms for patients. Introducing a clinical decision support system (CDSS) and integrating it into the clinical workflow and the electronic health record (EHR) can have a positive and significant impact on the prescription process by improving medication management safety. CDSSs are highly difficult to develop though and they often lack in usability, resulting in a loss of benefits a CDSS provides. However this issue could be lessened by utilizing user interface design principles during the CDSS development phase.

Objective: This study aims at validating the medication-related CDSS Sfinx and NjuRen using user interface design principles as a frame of reference.

Methods: A heuristic evaluation of both Sfinx and NjuRen was performed utilizing supplementary think-aloud sessions with subsequent interviews with physicians as a means to include clinical expertise.

Results: A list of ten violations against the user interface design principles were found and rated on their severity. Three additional problems were identified by the think-aloud participants that were not related to any user interface design principle.

Conclusion: Sfinx and NjuRen are quite well-designed in regards to their clinical content but lacking a bit as a computer system in general. This study could be an indication of the usefulness of applying user interface design principles as a means to increase a medication-related CDSS’s usability.

Keywords: Clinical decision support systems, User interface design, Medication alert systems, Health information technology, Design principles, Usability
Acknowledgements

First and foremost I would like to express my gratitude to my supervisor Sabine Koch who never stopped believing in me. Thank you for your support and keeping me on the right track along the way. Without your guidance there would not be a thesis at all.

Furthermore a big thanks to Emma Hultén at SLL. Thank you for taking the time to answer to all of my questions and help with the prototypes. Your immense commitment and involvement has been invaluable to me.

I must also express my deepest gratitude towards my family and friends for always being there. A special thanks to Kevin, my number one supporter.

Emma Molin
# Table of contents

Abbreviations .................................................................................................................. 7
List of figures ....................................................................................................................... 8
List of tables ......................................................................................................................... 8

1. Introduction ..................................................................................................................... 9
   1.1 The issue of medication errors ...................................................................................... 9
   1.2 Medication-related clinical decision support ............................................................. 9
      1.2.1 The usability of CDSSs ........................................................................................ 10
      1.2.2 Two examples of CDSSs - Sfinx and NjuRen ....................................................... 11
   1.3 User interface design principles ................................................................................. 16
      1.3.1 General usability heuristics .................................................................................. 16
      1.3.2 Design principles for medication-related CDSSs ................................................ 17
      1.3.3 Design principles for alerts in medication-related CDSSs .................................. 17
   1.4 Problem description ..................................................................................................... 19
   1.5 Aim and objectives ...................................................................................................... 19
   1.6 Research questions ..................................................................................................... 19
   1.7 Delimitations ............................................................................................................... 20

2. Methods .......................................................................................................................... 21
   2.1 Study design ................................................................................................................ 21
   2.2 Research methodology ............................................................................................... 23
   2.3 Data collection methods ............................................................................................. 23
      2.3.1 Meeting with developer ...................................................................................... 23
      2.3.2 Think-aloud with interview ................................................................................. 23
   2.4 Sampling ...................................................................................................................... 25
   2.5 Data analysis methods ............................................................................................... 26
   2.6 Ethical considerations ............................................................................................... 28

3. Results ............................................................................................................................. 30
   3.1 Heuristic evaluation of Sfinx and NjuRen ................................................................. 30
      3.2.1 Summary of violations ....................................................................................... 37
   3.3 Other results derived from think-aloud ..................................................................... 38

4. Discussion ........................................................................................................................ 40
4.1 Major findings..............................................................................................................40
  4.1.1 Recommendations for improvements ..............................................................41
  4.1.2 Previous research .........................................................................................42
4.2 Theoretical contribution ......................................................................................43
4.3 Strengths and limitations of the study ...............................................................43
  4.3.1 Strengths ......................................................................................................43
  4.3.2 Limitations .....................................................................................................44
4.4 Implications of findings .......................................................................................45
4.5 Suggestions for future research ........................................................................45
5. Conclusion ..............................................................................................................46

References ..................................................................................................................47

Appendices ..................................................................................................................50
  Appendix A – Consent form.....................................................................................50
  Appendix B – Think-aloud.......................................................................................51
  Appendix C – Interview questions ........................................................................52
  Appendix D – Additional screenshots of Sfinx and NjuRen..................................53
Abbreviations

CDSS – Clinical decision support system
DDI – Drug-drug interaction
eGFR – Estimated glomerular filtration rate
EHR – Electronic health record
SLL – Stockholms läns landsting/Stockholm county council
List of figures

Figure 1: The Janus toolbar as it appears in the EHR .............................................11
Figure 2: Screenshot of Sfinx accessed through the web ............................................12
Figure 3: Screenshot of Sfinx accessed through Janus toolbar in EHR .........................13
Figure 4: Screenshot of an early prototype of NjuRen accessed through the EHR ..........15
Figure 5: Illustration of the research process of this study .........................................21
Figure 6: Screenshot of Dedoose displaying codes, excerpts and documents ..............27
Figure 7: Screenshot of NjuRen used in think-aloud sessions ....................................53
Figure 8: The window in NjuRen where it is possible for the user to enter values and calculate their own absolute eGFR .................................................................54
Figure 9: Classification of DDIs in the system documentation ....................................54
Figure 10: Doing a search for drug or substance in Sfinx accessed through the web ....55
Figure 11: Example of a severe nonintrusive alert in Sfinx ........................................55

List of tables

Table 1: Results from evaluation using general usability heuristics .........................33
Table 2: Results from evaluation using design principles for medication-related CDSSs.................................................................34
Table 3: Results from evaluation using design principles for alerts in medication-related CDSSs.........................................................................................................................37
Table 4: List of all direct violations against user interface design principles ..............38
1. Introduction

1.1 The issue of medication errors

The process of prescribing medicine to patients has been recognized as an international issue (1) due to the process being complex and error prone, causing preventable harms for patients (2). Among all medical errors occurring in health care systems today, medication errors are the leading cause of these and constitute the majority (3,4). Furthermore medication errors are prevalent through all stages of medication-use in health care processes but most commonly they occur at the point of prescribing and administration (4), relating to the complexity of the processes. In order to understand what constitutes a medication error, being a type of medical error, one can review definitions such as one provided by the National Coordinating Council for Medication Error Reporting and Prevention in the United States (5):

“A medication error is any preventable event that may cause or lead to inappropriate medication use or patient harm while the medication is in the control of the health care professional, patient, or consumer. Such events may be related to professional practice, health care products, procedures, and systems, including prescribing, order communication, product labeling, packaging, and nomenclature, compounding, dispensing, distribution, administration, education, monitoring, and use.”

A primary cause of general medical error rates and their increase is the complexity and intricacy of today’s health care systems (6). The system governs what the people in them do thereby creating an environment where errors usually lie outside the control of individuals (4). This means that little blame can be put on the individuals and one must look beyond in order to solve the underlying system failures that create situations under which medication errors occur (4). These situations could for instance be solved by designing health care systems that enables the detection of errors before considerable damage occurs (6). This is significant to acknowledge since there is a need of lowering the occurrence of medication errors in order to enhance patient safety and improve health care systems.

1.2 Medication-related clinical decision support

In response to the increasing number of medication errors a proposed and recommended intervention is to introduce a computerized Clinical Decision Support System (CDSS). In recent years it has been established that incorporating a CDSS into clinical workflow and integrating it with the electronic health records (EHR) can have a positive and significant impact on the prescription process by improving medication management safety (2,7,8). The aim of clinical decision support is to aid the clinicians with diagnostic and other care-related decisions by providing relevant and usually patient-specific information at various points in the care process (9).
A CDSS is more specifically software consisting on a basic level of three main components; (a) an interface somehow integrated or connected to an EHR in order to access relevant patient data, (b) a knowledge base, and (c) a dedicated inference engine that will match the retrieved patient data with suitable medical knowledge with the purpose of providing a recommended care plan to clinicians for consideration (3). Thus in relation to medication prescribing a CDSS could for instance check drug-drug interaction. Meaning it will access drugs that are intended to be prescribed as well as drugs already prescribed in the EHR and find out, by using the knowledge base to as reference, how these drugs will interact. The result is thereafter a recommendation based on what was discovered. For instance the CDSS could recommend the physician to change or remove a drug since it appears to interact badly with another and this could potentially harm the patient.

Apart from drug-drug interaction, which is one of the simpler functionalities, a medication-related CDSS can have other basic functionalities as well as other more advanced ones. What differentiate these is that the medical issues related to the basic functionalities are more straightforward. Some examples of the basic functionalities are; drug dosing guidance, drug-allergy checking and, duplicate therapy checking meaning trying to stop prescriptions of different drugs with similar effects. Some examples of the more advanced functionalities are; dosing support for renal insufficiency, drug-pregnancy checking and, medication-related laboratory testing meaning that some drugs require laboratory tests to be done prior to administration and regularly during the treatment. These are the most common features that a medication-related CDSS could potentially have and it gives an idea of the potential it has to reduce medication errors. (2)

1.2.1 The usability of CDSSs

Even though introducing a CDSS has the potential of bringing many benefits to clinical practice, a highly effective CDSS is challenging to both develop and implement (8,10). One of the core issues with CDSSs is that they often face acceptance problems among the users and poor usability appears to be the leading barrier for adopting CDSSs and a deterrent for its routine use (7,9,10). Thus poor usability causes rejection of the system resulting in a loss of potential benefits for the prescription process. In addition poor usability could cause medication errors as well, thus also affecting the patient negatively and causing more problems than those already present (11). One reason for the common problem of poor usability in CDSSs is that there are known design principles but these are missing in today’s health information technology (1), which comprises CDSSs. Another reason is that there are no standardized, agreed upon best practices and design principles which in return causes CDSS vendors to use their own guidelines and standards for design (10).
1.2.2 Two examples of CDSSs - Sfinx and NjuRen

In Stockholm County, Sweden a medication-related CDSS called Janus toolbar (Janusfönster in Swedish) has been made available to public care units as well as some others around Sweden. The Janus toolbar has been developed by the Stockholm county council (Stockholms Läns Landsting, SLL) and is integrated with and made accessible from the EHR first and foremost. In the EHR the Janus toolbar appears as a bar with six different buttons (see figure 1) and each button leads to a specific medication-related CDSS function such as; drug-drug interaction, drug-pregnancy checking, drug-breast feeding checking, information about drug side effects, prescription support for renal insufficiency and, elderly-drug checking. (12)

The different buttons act as warnings by changing color and indicating a possible problem based upon patient-specific information retrieved from the EHR. These warnings are all classified based on their severity and activates when the user has opened the patient’s medication list in the EHR. Also when the user creates a new prescription the color and information on the buttons will update and change depending on the new prescription. (13)

![Figure 1: The Janus toolbar as it appears in the EHR](12)

Recently two of the functionalities in the Janus toolbar have been under development by the E-health and strategic IT (E-hälsa och strategisk IT) division at SLL. One of the functionalities is the drug-drug interaction (DDI) which have been updated and released already. The other functionality is the prescription support for renal insufficiency which is currently being updated and improved with an expected release in 2017.

The DDI functionality is called Sfinx (Swedish-Finnish Interaction X-referencing) and provides the prescriber with short assessments of all interactions between drugs including medical consequences and recommended actions for instance. Sfinx covers all authorized substances in Sweden and/or Finland as well as some of the interactions drugs can have with alcohol, food and smoking for instance. The aim of Sfinx is to include all clinically relevant drug interactions documented in scientific literature, mentioned in the summary of product characteristics and interactions that are expected based on the medicines’ circulation and excretion. (14)

Sfinx can either be accessed through the web by anyone or through the Janus toolbar in the EHR by a clinician. Sfinx will function slightly different depending on how you access it. If it is accessed through the web by anyone it is possible to find DDIs by searching for substances or brand names (see figure 2), meaning that you have access to the entire knowledge base in the system. If Sfinx is on the other hand accessed through the Janus toolbar in the EHR it automatically retrieves drugs present in the patient’s EHR and presents how those interact
(see figure 3). The user is not able to search for any other substances or drugs in Sfinx at that point.
The prescription support for renal insufficiency functionality in Janus toolbar is called NjuRen (njuren means ‘the kidney’ in Swedish) and is one of the newest additions to the Janus toolbar. The purpose of NjuRen is to provide the prescriber with evidence-based recommendations about dosing of drugs based on the patient’s renal function. What NjuRen does today is to calculate and display a relative eGFR (estimated glomerular filtration rate) based on serum creatinine, age and gender of the patient which is retrieved from the EHR. The relative eGFR, which indicates the renal function, then dictates the recommendations for
the patient’s current medications. NjuRen does also allow for input of height and weight in order to calculate an absolute eGFR. In terms of access, NjuRen can only be accessed either by the Janus toolbar in the EHR or through the web on SLLnet which is a sort of intranet. (15)

In regards to the version of NjuRen under development, displayed as figure 4, it is intended that the application will be able to extract length, weight and cystatine c values as well as serum creatinine, age and gender of the patient. The outcome of this update is the possibility to provide an enhanced estimate of the eGFR.

As for previous scientific studies in regards to Sfinx or NjuRen there has been for instance an evaluation done on usage patterns and user perceptions of Sfinx by Andersson et al. (16). This study sent out a questionnaire to all registered users of the web application of Sfinx with the purpose of gaining an insight to how prescribers and pharmacists perceive and utilize Sfinx in their clinical work. The results showed that most users used the system during a patient consultation and that most prescribers and pharmacists changed their handling of patients based on the information received by Sfinx.
Figure 4: Screenshot of an early prototype of NjuRen accessed through the EHR
1.3 User interface design principles

The user interface is what constitutes the computer system by most users because it is what you see when using it, thus making it the most important part of all systems. The user interface can, depending on a poor or good design, affect the users’ experience and utilization of the system. Thus the design of the user interface is something that the users value highly. For instance a poorly designed user interface can appear confusing for the user, resulting in great difficulties which in turn lead to mistakes, workarounds or complete abandonment of the system. On the other hand a good user interface design can result in great pay-offs such as making tasks more efficient and for instance by formatting inquiry screens based on respectable design principles it is possible to reduce decision-making time. (17)

Applying user interface design principles is a means to obtain a well-designed user interface for any type of system (17) and it also applies to medication-related CDSS. Following appropriate user interface design principles is essential when developing highly effective CDSSs (9). When it comes to choosing which design principles to apply, there is an abundance of different design principles that are applicable to all types of computer systems. One of such is Nielsen’s 10 usability heuristics for user interface design (18) which are quite well-known and widely applied on computer systems in order to increase usability. However computer systems differ greatly and a medication-related CDSS cannot be designed in the exact same way as for instance one intended for banking or retail. Thus attempts have been made to gather and compile user interface design principles specific for medication-related CDSSs such as work done by Horsky et al. (9,19) for instance.

In the following sections a few of both general usability heuristics as well as more specific for medication-related CDSSs are presented. The user interface design principles has been divided into three different levels of specificity in regards to medication-related CDSS in order to cover a wide spectrum of design principles applicable to this kind of system and user interface.

1.3.1 General usability heuristics

Jakob Nielsen has developed a set of ten general usability principles or heuristics for user interface design based on a factor analysis of a number of usability problems. The first list of principles was developed in 1990 by Nielsen and Rolf Molich but was revised until the final version which was established and published by Nielsen in 1994 (18). Still today these are widely taught and applied in user interface design. The ten usability principles are as follows (18):

- Visibility of system status
- Match between system and the real world
- User control and freedom
- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognize, diagnose, and recover from errors
- Help and documentation

1.3.2 Design principles for medication-related CDSSs

Horsky et al. (9) have recognized both the problem with poor usability and the lack of standards of design when it comes to CDSSs ready to be adopted. Therefore in an attempt to provide developers with some good design principles the authors have done a targeted review of available best practices, procedures and lessons learned and compiled them in a short compendium. The purpose of their review is to clarify design goals for medication-related CDSSs. But also to be a first step towards recognizing the need for larger national or international initiative of establishing standardized guidelines for user interface design of CDSSs in order to advance within the field. The result of Horsky et al.’s (9) research was a list of 15 recommendations specifically for medication-related CDSSs. Some of these were for instance:

- Consistent terminology
- Format text to visually associate drug categories
- Show relevant patient-specific information
- Concise language
- Emphasize differences in similar drug names
- Visually distinct screens for confusable items

1.3.3 Design principles for alerts in medication-related CDSSs

A major part of medication-related CDSSs is their alerts, both interruptive and non-interruptive and has thus been sometimes given a separate set of design recommendations. Interruptive alerts are those that disturb the workflow and demand a response or an action by the clinician before proceeding. For example it could be a so called pop-up with the warning on which you actively have to do something with in order to proceed. These interruptive alerts for the more serious warnings and their purpose are to force the clinician to recognize a possible hazardous situation. Non-interruptive alerts are for the less serious warnings and they do not disturb the clinical workflow due to them appearing as info-buttons or links on which you can choose to open or not. It is preferable today to have CDSSs that eliminates the need for interruptive alerts by guiding the user instead, since too many interruptions can cause irritation amongst the users. (10)
Payne et al. (10) are one group of researchers that has focused on the user interface design of both interruptive and non-interruptive alerts. More specifically they have through conducting a series of meetings with a group of experts consisting of people with a background in either medicine, informatics or computer interface design looked at the design of drug-drug interaction (DDI) alerts. A result of their study was for instance a list of components that a DDI alert should include:

- Drugs involved
- Signal word indicating the level of seriousness
- Clinical consequences
- Mechanism of the interaction
- Contextual information/modifying factors
- Recommended action(s)
- Evidence

Horsky et al. (9) have aside from their list of general user interface design recommendations also provided specific suggestions for alerts and reminders. These recommendations were compiled in the same targeted review described previously. Some of the recommendations they provide specifically for alerts are:

- Interruptive alerts should be reserved for the two or three highest levels of severity
- Content of the alert should be limited to 1-2 lines with a justification separated by white space.
- Alert prioritization
- Meaningful color coding
- Revise trigger rules

Furthermore design recommendations have been compiled exclusively for interruptive alerts. Horsky et al. (19) have done a targeted literature review of lessons learned during CDSS implementations. The lessons learned from success and failures were then interpreted in relation to human-computer interaction and software usability principles. The result of the work done by Horsky et al. is a list of optimal design attributes of interruptive alerts and some of these were as follows (19):

- Follow common design conventions
- Emphasize severity level
- Include contextual information
- Include triggering rule
- Concise directions
- Labels (static information) are deemphasized
- Allow entering of override reasons
1.4 Problem description

Several CDSSs today suffer from poor usability due to the serious challenge of developing an effective CDSS (7) and potentially due to the lack of using user interface design principles in the development process (1). The resulting poor usability causes rejection of the system and has become a leading barrier to CDSS adoption (7,9,10). By rejecting the system all potential benefits that a CDSS could have provided are lost. Additionally a poorly designed user interface can result in mistakes and issues that ultimately could cause medical errors and affect the patient safety negatively (11). Thus it becomes of the utmost importance to enhance usability of CDSSs and one way to do that is to utilize user interface design principles during the development phase. However there is little research done on how one could apply the design recommendations, probably because there are no standardized and agreed upon design conventions to follow and the ones available are not older than circa five years. This study will contribute to this research field by providing an example of how a medication-related CDSS could be analyzed and validated using user interface design principles.

1.5 Aim and objectives

This study aims at validating the medication-related CDSS Sfinx and NjuRen using user interface design principles as a frame of reference. In order to fulfill the aim, the study makes use of a number of objectives that would essentially guide the work towards the main aim. The first objective is to determine a base of user interface design principles, both generic and specific for medication-related CDSSs. The second objective is to perform a heuristic evaluation and validate how well the interface design of Sfinx and NjuRen fit the compiled design principles gathered. Finally the third objective is to determine areas where Sfinx and NjuRen fell short in the evaluation and provide some suggestions for improvements based on the design principles in order to help the responsible developers in their future development of the systems.

1.6 Research questions

By evaluating the two CDSSs the study intends to answer the following main research question:

- How well does the user interface design of medication-related CDSSs Sfinx and NjuRen live up to both generic and specific user interface design principles?

In addition, the main research question is accompanied by a sub-question that would be answered as well along the way in order to fulfill the main aim of the study:

- Are there any areas for improvements in Sfinx or NjuRen based on the user interface design principles and if so, what do the principles propose to be done?
1.7 Delimitations

This study evaluates only Sfinx and NjuRen and not the Janus toolbar or the other decision support functionalities attached to it. The study would have become too extensive and time-consuming. In addition it appeared as a justifiable delimitation since the evaluation regards newly developed versions of Sfinx and NjuRen that now differs in design from the other decision support tools. By evaluating the new designs this study could provide some support for improvements on not only Sfinx and NjuRen but also if the other decision support tools are to be redesigned and updated as well. Furthermore since the prototype of NjuRen was not finished during the time-period of this study, it was only possible to evaluate NjuRen accessed through the EHR and not via the intranet.
2. Methods

Based on how the aim, objectives and research problem was outlined the research design in this study was a qualitative heuristic evaluation complemented with three other data collection methods. The study followed four basic steps illustrated in figure 5 and was initiated by a literature review where a list of user interface design principles and recommendations was established. A subsequent step was to review the list of principles and determine the methods needed in order to be able to validate how well the CDSSs correlated to the principle. The outcome of this revision was three different procedures; asking the developers, think-aloud with semi-structured interviews and, validation done by the researcher). Consequently the third major step was to collect data through think-aloud sessions with subsequent interviews with physicians operating within Stockholm County and meeting the developers. Finally with the addition of collected and analyzed data from think-aloud sessions and the developers a heuristic evaluation was performed by validating how well system properties of NjuRen and Sfinx lived up to the compiled list of design recommendations.

![Figure 5: Illustration of the research process of this study](image)

2.1 Study design

It is the aim of the study that determines the appropriate research design (20). Thus depending on how the research problem and the study’s purpose are framed it determines which method is the most fitting. This study had the aim of validating Sfinx and NjuRen using user interface design principles as a frame of reference. Based on the aim this study took the shape of an evaluation. Ammenwerth et al. (21) defines evaluation research as:

“...the act of measuring or exploring properties of a health information system (in planning, development, implementation, or operation), the result of which informs a decision to be made concerning that system in a specific context”.

Thus evaluation allows for the exploration for properties of a health information system, under which CDSSs falls under, making it an appropriate research design for this study. With evaluation as a basis for the study’s design, Brender McNair (22) further elaborates on concepts within the field of evaluation studies within health informatics. One term which falls
under the concept of assessment and evaluation is validation. Validation is defined by Brender McNair (22) as the:

“Act of comparing properties of an object with the stated goal as a frame of reference”.

Thereby validation became an appropriate fundamental concept for this evaluation where the objects are Sfinx and NjuRen, properties are the user interface and the stated goal as a frame of reference is the gathered design recommendations.

Furthermore there are several types of evaluation methods to choose from and an analytical approach was chosen initially based on the study’s aim. Analytical approaches consist of methods relying on analysts’ judgments and their analytic techniques used when performing evaluations on user interfaces. These types of methods seldom involve the contribution of users. Under the umbrella of analytical approaches there are amongst others inspection-based evaluations which entail at least one expert acting as user with the intention of identifying possible usability problems with a computer system. This could be done both on complete systems or interfaces as well as prototypes or beta versions and it is a cost-effective way to identify a system’s shortcomings. (23)

One of the most common types of inspection-based evaluations is the heuristic evaluation which became a general framework for this study due to its suitability to the aim. The heuristic evaluation was founded by Jakob Nielsen with colleagues and it entails at least one usability expert evaluating a user interface by attempting to identify possible violations to a set of selected heuristics. The evaluation follows these three steps: 1) Selection of relevant heuristics based on the type of system to be evaluated, 2) Evaluation of user interface against selected heuristics by experts and, 3) Identification and rating the severity of all violations found. The rating is done by scoring each violation on a scale of 1-5, where 1 indicates a cosmetic problem and 5 signifies a catastrophic problem. Usually the heuristic evaluations utilize usability experts but it has happened that they have been replaced with subject matter experts instead. (23)

Thus in conclusion the chosen study design was heuristic evaluation since that specific method was closely related to the aim of the study which was to primarily without the involvement of users validate Sfinx and NjuRen using user interface design principles as a frame of reference. The researcher took the role of the evaluator but due to the nature of medication-related CDSSs it became significant to involve a small number of users to act as subject matter experts as well in order to add clinical expertise. By including subject matter experts a more complete heuristic evaluation of the CDSSs was possible. However the users were not involved in the actual evaluation, their role was to provide input where the evaluation otherwise fell short.
2.2 Research methodology

Besides research design there are two main divisions of research methodology and those are qualitative and quantitative research methods. Quantitative research in general terms focus on the collection of numerical data and quantification of different aspects of social life. This type of research is related to natural science and an objective perception of reality. Qualitative research on the other hand is more interested with words rather than numbers. It has an interpretive approach and collects data by observing how individuals perceive their reality. Instead of having an objective perception of reality, qualitative research prefers to view reality as an ever changing attribute which belongs to individuals’ constructive capability. These are the most basic distinctions between quantitative and qualitative research and even though they differ it is possible to combine them and use research methods from both, which is referred to as a mixed methods approach. (24)

Evaluation can take the shape of both quantitative and qualitative research. Which method it is going to be is determined by whether it is possible to somehow quantify and measure the results or whether the result is going to be entirely descriptive (22). Based on this the study applied a qualitative approach due to the frame of reference being descriptive recommendations which have to be interpreted and applied.

2.3 Data collection methods

As described and illustrated in figure 5 there were two main data collection methods acting as a complement for the subsequent heuristic evaluation. One of the data collection methods was conversation with developer and the other think-aloud sessions with semi-structured interviews.

2.3.1 Meeting with developer

In order to find out things about Sfinx and NjuRen unseen to the user an informal meeting was held with a contact at SLL deeply involved with the development of both systems. This meeting allowed for questions such as if there are any possibilities to customize the contents of Sfinx and NjuRen for instance. Another benefit was the possibility for the researcher to gain insights to how NjuRen is intended to function in the future if the development proceeds as expected. This allowed for a more correct evaluation of NjuRen even though it was far from finished.

2.3.2 Think-aloud with interview

For this study two methods of exploring the user’s interaction with the user interface were chosen, think-aloud and semi-structured interview. The rationale of adding such methods to
this assessment was to capture aspects of the CDSSs which required the user to have clinical knowledge and expertise. For instance it is difficult to determine if the language used in the CDSS is appropriate with phrases and concepts used in the real world, the clinical setting, if one does not possess clinical expertise. Thereby physicians were chosen for these examinations since they constitute the main targeted group of users of Sfinx and NjuRen.

There are several basic techniques for examining how users interact with the computerized system such as: think-aloud, interview, questionnaire, observation, automatic logging of cursor movements and so on (25). Specifically for this study it was significant to know what the user thought of the CDSS in order to determine their opinions and rationale for interacting the way they do. Techniques such as observation or automatic logging of cursor movements would not have brought any value to the study since they lack the vocalization of the participant’s opinions. Other techniques such as questionnaire and interview bring a retrospective report which is flawed since things may be forgotten and lost easily (26). A think-aloud on the other hand capture reasoning and the thought-process as it happens making it more valuable and complete than recollections (26). Thus the think-aloud appeared as the most appropriate technique for this study.

However using only one technique is often viewed as insufficient and it is therefore common to combine several of the techniques mentioned earlier (25). Fonteyn et al. (26) suggests combining the think-aloud with a follow-up interview in order to obtain the fullest description possible of the participant’s reasoning. Nielsen et al. (25) also acknowledge that some studies use something called a think-after where the researcher is allowed to interview and ask questions about the think-aloud session. This was utilized in this study in order to truly capture all aspects of the user interface design principles where clinical expertise was necessary. The reason for disregarding a questionnaire was to still have the ability to ask questions specifically about things that happened during the think-aloud, making an interview more appropriate.

A think-aloud session consists of asking people to, while performing a task, verbalize their thoughts. The researcher’s role is to solely observe the participant and typically audiotape or videotape the session. The main benefit of the think-aloud technique in comparison to others is that it reveals the thinking process in relation to the participant’s concurrent perceptions. In addition it has been shown that despite having a small number of participants, the think-aloud technique stills provide valuable and rich data for analysis. The think-aloud technique has however been criticized due to the capacity of the working memory while thinking aloud. Some researches argue that this issue would hinder the cognitive process of the participant thus making it more difficult than usual to perform a task demanding a high cognitive load. (27)

An interview is good for gaining insights in complex situations and peoples’ feelings, opinions, experiences and emotions. There are three types of interview; structured, semi-structured and unstructured. The structured interview lends itself for some kind of standardization where the researcher asks identical questions to each respondent. The semi-
structured interview does also contain a predetermined set of questions but the researcher is allowed to be more flexible with the questions. The main purpose of the semi-structured interview is to allow for the respondent to talk more freely and elaborate on things that comes to mind during the interview but still keeping some control of the interview to the researcher. Lastly unstructured interviews are mostly focused on what the respondent wants to talk about. The researcher’s role is to get the ball rolling by introducing a topic and then being as unobtrusive as possible. Semi-structured interview was chosen for this study in order to stay on topic and keep the freedom of changing the questions slightly depending on what was said during the think-aloud. (28)

The think-aloud sessions with the physicians were scheduled at different times between the dates April 5 – April 21. The sessions were performed in Swedish since it is the native tongue of both participants and researcher. By allowing the participant to speak in their native tongue the cognitive load is reduced since they do not have to translate as well as thinking about the task. The result is a more accurate outcome of what the participant was thinking.

Each session began with an introduction of the setup of the session and the signing of a consent sheet (see appendix A). The consent sheet was developed using guidelines and a template from Denscombe (28) including aspects such as participation being voluntary, information about the study and confidentiality for instance. After making sure that the setup was clear the think-aloud could begin. The participant was given a hypothetical patient case (29) which had been augmented slightly by changing two of the drugs in order to increase the variation amongst the drug interactions. The case was followed by a number of tasks to perform in both Sfinx and NjuRen (see appendix B). Sfinx had been released at the time of the think-aloud. NjuRen however was far from finished which meant that the prototype was flawed and not functioning properly. In order to compensate a screenshot was included created by SLL depicting how it will somewhat appear in the future (see figure 7 in appendix D). As a complement to this static screenshot the participants were able to also see the flawed functioning prototype in order to click and explore more (figure 4 in the introduction). However it was clearly stated in the think-aloud instructions that they should have an open mind since many functionalities does not work as intended.

Furthermore the tasks were kept simple with the only purpose of familiarizing the participant to the system. At the same time the tasks were kept open and not to specified in order to allow the participant to steer their focus on what captures their interest. After the think-aloud a short interview was conducted with a small number of questions (see appendix C). The reason for keeping the think-aloud and interview short and concise was the limited time available for the participants to participate.

2.4 Sampling

According to Nielsen (30) one should plan for including $4 \pm 1$ subjects in a think-aloud study. This study aimed at having five participants but ended up with four due to difficulties
reaching participants and time restrictions. However it is still within the accepted spectrum according to Nielsen’s recommendations. A small sample is fundamental in think-aloud studies since it seeks exhaustive rich data (26).

In terms of sampling this study employ an exploratory sample. An exploratory sample lends itself well to qualitative studies with the purpose of exploring relatively unexplored areas and gaining new insights. The specific approach chosen was non-probability, meaning that the selection of the sample was not randomized but included an element of choice in the selection process. The reason for choosing a non-probability approach was that the goal of this study is not to be representative and some influence was needed on the selection of participants. (28)

Under the umbrella of the non-probability approach there are several sampling techniques available. A technique called purposive sampling, meaning hand-picked for the topic, was primarily used. Purposive sampling is suited for the acquisition of an exploratory sample and can be used as a way of obtaining high quality information by selecting people likely to have expertise or experience of the research topic (28). This sampling technique was valuable to this study since it required the expertise of an actual user, a physician. In addition in order to save time in the think-aloud session it was decided to only include physicians with at least some experience with Sfinx and NjuRen. Thus reducing the time spent explaining the CDSSs and how they function in general. The purposive sampling was performed by contacting public clinics around the Stockholm County, thereby ensuring that the clinic at the very least have access to the CDSSs. Contact was made through sending e-mails and explaining what was required of the physician if they were to participate. The e-mail was sent to the operational manager at each clinic who then referred to interested physicians. Thus including some aspects of the non-probability technique called snowball sampling when the researcher is referred by one person to the next and in so doing builds on their sample (28).

2.5 Data analysis methods

The first step of analyzing data from think-aloud sessions is to transcribe the audio recordings (26), this is also the case for interviews (28). Transcription is the process of writing down in sentences what was said during an audio-recorded interview for instance, it does not have to be word by word depending on the use of the data collected (28). The transcripts made for the four think-aloud sessions with following interview were written almost word by word. Some information was left out if it was purely related to the think-aloud session itself and not its contents. Meaning that information regarding instructions and questions not related to the CDSSs themselves was removed due to its irrelevance to the study’s purpose. In addition, as previously mentioned, the think-aloud sessions were performed in Swedish thus they were also transcribed and analyzed in Swedish. By doing so little is lost in translation and it is only when transcripts are referred to in the following results section as they were translated into English.
The next step is the encoding of the transcripts. This is done by using predetermined coding categories to label or code an excerpt of the text. Before doing so the length of the excerpts should be defined and for this study it was decided to have each sentence being an independent excerpt. This allowed for keeping some of the meaning and context since each segment is to be treated separately of the surrounding text. (25)

For the actual coding this study made use of computer assisted qualitative data analysis software. The chosen software was the web-based tool called Dedoose (see figure 6) which is free of use for 30 days. This type of software aids in the analysis of qualitative data by providing for instance secure storage of data, coding functionalities and simple retrieval of data (28). Consequently when the transcribing process was finished, each document was uploaded into Dedoose. Dedoose then allows you to create coding categories and sub-categories which you use to label excerpts and thus code the transcripts. By coding the text, Dedoose then offers functionalities that support qualitative analysis such as for instance code filtering, code frequency and code co-occurrence etc.

The codes used were the design principles found in the literature review. The purpose was to ease the comparison and validation of each design principle. By having the used codes becoming a list of design principles mentioned by the participants it simplified the succeeding heuristic evaluation process by clarifying where the participants had input for the evaluator to consider. Furthermore another code called ‘Other’ was also included in order to allow for a categorization of extra commentaries not related to the chosen design principles.

Finally all of the collected and processed data from transcripts, notes from meetings with developers as well as notes taken during the think-aloud sessions were then utilized in the heuristic evaluation as a support to the researcher as evaluator. This refers to the final step of a heuristic evaluation where violations against the frame of reference are identified by the
evaluator and scored based on their severity. The actual scoring was performed using the scale of 1-5 as previously mentioned in the assessment of each violation. The severity was based on three factors by Nielsen (18):

- The frequency with which the problem occurs: Is it common or rare?
- The impact of the problem if it occurs: Will it be easy or difficult for the users to overcome?
- The persistence of the problem: Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?

These three factors combined determined one severity score in total in order to allow for an overall assessment that facilitates prioritization.

2.6 Ethical considerations

According to Bryman (24) there are fundamental ethical questions in qualitative research concerning; confidentiality, anonymity, integrity, voluntariness and consent requirement. To start with, confidentiality concerns how the data, especially personal data, is handled and protected (24). For this study confidentiality have primarily been achieved through never having personal information connected to the respondent’s answers, thus if access is unfortunately gained to the data it is not possible to deduce who the respondent is. Another fundamental ethical concern of research is anonymity. The study had little ethical concerns regarding anonymity and it ties into what was previously stated, that audio-files and notes were anonymized immediately and only connected to a given ID. It was also made sure that the anonymity and confidentiality was clearly stated in the consent form (appendix A).

Furthermore personal integrity is another ethical issue in research and it concerns the protection of the respondent and to not intrude on their privacy (24). Due to the high level of anonymity, the respondents’ privacy was protected. Additionally this study did not target a vulnerable population and the topic did not concern anything sensitive which reduced the intrusion of the respondent’s privacy. The fourth ethical conundrum is voluntariness. This has been considered for this study by explicitly stating in the consent form that participation was voluntary and the participant had the option to end the think-aloud session at any time.

As for the consent, it is the fifth and final ethical issue. This principle has to do with that those who participate in a study should be fully informed about the purpose and structure of the research (24). A consent form, as previously mentioned, with general information about the study was presented prior to the think-aloud session and it was mandatory to sign it in order to participate. Thus no data was collected and shared without the consent of the participants.

Lastly this study was done in collaboration with the division of E-health and Strategic IT at SLL since Sfinx and NjuRen belongs to them. However this study was performed independently with merely support from contacts at SLL. Support in terms of answering
questions about the systems and assuring access to the prototypes. Moreover no financial aids were given by any organization, hereby further emphasizing this research as a stand-alone study with no conflict of interest.
3. Results

This section presents the results and findings that derived from this study and its different methods. The study was initiated by a literature search which resulted in four different sources of user interface design principles described briefly in the introduction. For the evaluation only three of these were utilized since the ones by Horsky et al. (19) were only applicable to interruptive alerts and it turned out that neither Sfinx nor NjuRen had any interruptive alerts. Further on in this chapter the results from the heuristic evaluation are presented where each and every design principle from the three chosen sources is described and how it compared to Sfinx and NjuRen. This section includes data recorded from think-aloud sessions and conversation with developers. Lastly a review of other results generated from the think-aloud is presented in order to describe the general outcome of the session.

3.1 Heuristic evaluation of Sfinx and NjuRen

The main part of this study was the heuristic evaluation of Sfinx and NjuRen using the predetermined user interface design principles as a frame of reference. As previously mentioned most of the evaluation was done by the researcher taking the role as evaluator using data collected from think-aloud sessions as a supplement where clinical expertise was necessary. The results from the heuristic evaluation have been put in three tables in order to present the outcome in a structured way. Each table represents the results for the three different levels presented in the introduction starting with the general usability heuristics. For each user interface design principle a result has been recorded as well as a severity score of a potential violation. If there is no violation against the design principle or if it was not applicable, no score has been given.

Furthermore there is a column called ‘Method’ which indicates if the specific result was based on data collected from either the think-aloud sessions or conversation with developers. Abbreviations have been used for this column where P indicates think-aloud participants and D – information provided by developers. If no method is stated the result is only based on evaluation performed by the researcher.

All in all 38 user interface design principles were utilized for the heuristic evaluation. The list has shrunk in this compilation due to some overlapping where two sources have had the same design principle. This was only the case in the category called ‘Design principles for alerts in medication-related CDSSs’ presented in table 3.

Finally in the following tables there are references to screenshots from both Sfinx and NjuRen in order to clarify some results. All of the figures referred to can be found in appendix D. In addition it is also significant to point out that NjuRen was evaluated as if buttons and eGFR calculations were functioning as intended.
<table>
<thead>
<tr>
<th>No</th>
<th>User interface design principle</th>
<th>Results from validation</th>
<th>Method</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Visibility of system status: The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.</td>
<td>Well done in NjuRen where it is clearly stated what the measurements taken from the EHR are. The user does not have to guess what has been accounted for in the calculation. In Sfinx if a commentary is sent there is a message indicating that it has been sent. However this message disappears almost immediately, giving the user almost no time to read it. Otherwise good visibility of system status in Sfinx.</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Match between system and the real world: The system should speak the users’ language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.</td>
<td>None of the participants had any complaints about the language. Participant 3 said “That was easy and this is similar to reality”. Concepts appeared familiar to the users.</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>User control and freedom: Users often choose system functions by mistake and will need a clearly marked &quot;emergency exit&quot; to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.</td>
<td>Since the purpose of a CDSS is not to control the user but to aid in the decision making process, no actual decision making can be done in the system. Thus Sfinx and NjuRen have little functionalities where you could actually do something. NjuRen does however have functionality for calculating your own eGFR and there is an option to cancel and not use this calculated eGFR in case a mistake was made or other. Thus undo is supported in this one case.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Consistency and standards: Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.</td>
<td>Sfinx and NjuRen were consistent in language and color coding. Participants recognized it and felt comfortable especially with the categorization of interactions. However in NjuRen it appeared disconcerting to one participant that the complete recommendation did not include both name of the product and the substance as they are used to and as it did in Sfinx. In addition it was recognized by another participant that interactions in Sfinx were classified as D3, D2, D1 etc. But in NjuRen no such numbering of A-D was included.</td>
<td>P</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Error prevention: Even better than good error messages is a careful design which prevents a problem from occurring in the first place.</td>
<td>This could be done better in NjuRen where the user is allowed to enter any value, even a negative one, in the eGFR calculator. (see figure 8) This could easily be prevented by not allowing any unrealistic input. Furthermore both systems include warnings</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>No</td>
<td>User interface design principle</td>
<td>Results from validation</td>
<td>Method</td>
<td>Score</td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------</td>
<td>-------------------------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>on how to interpret the presented information so that no decisions are made based on non-applicable information. That is their type of error prevention.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Recognition rather than recall: Make objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.</td>
<td>Participants agreed upon that the system was easy to navigate and by observing them they appeared to almost instantly know where to click. Instructions for use of the system were easy to find in the web-based Sfinx accessible to anyone but somewhat hidden in the others. There was no clear button of where to find the documentation in neither Sfinx nor NjuRen. In addition since personal data is included from the EHR, the user never has to remember what was written in the EHR which is a good thing.</td>
<td>P</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>Flexibility and efficiency of use: Accelerators-unseen by the novice user may often speed up the interaction for the expert user to such an extent that the system can cater to both inexperienced and experienced users.</td>
<td>The systems do not have any accelerators for the expert user.</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Aesthetic and minimalist design: Dialogues should not contain information which is irrelevant or rarely needed.</td>
<td>All participants agreed upon that nothing in the systems seemed irrelevant. Information that seemed less relevant could be accessed via links instead of taking up space, competing with the more relevant information.</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Help users recognize, diagnose, and recover from errors: Error messages should be expressed in plain language, precisely indicate the problem, and constructively suggest a solution.</td>
<td>During the evaluation it was not possible to find an error message. This is somewhat due to what was mentioned in design principle 3. However in figure 8 it is clear that faulty parameters had been entered. Instead of writing NaN, which may be incomprehensible to some, there should be a better explanation of what went wrong. It is understandable that something is wrong but it should be expressed better.</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Help and documentation: Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete</td>
<td>A major problem is that no help or documentation can be found in NjuRen currently. In Sfinx it can be found, even though it is not easy (see design principle 6). Locating issues aside the system documentation for Sfinx is well-structured, informative and no too overwhelming. In Sfinx the system is also helpful in terms of directing the user to specific instructions in the documentation. For instance by clicking</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>User interface design principle</td>
<td>Results from validation</td>
<td>Method</td>
<td>Score</td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------</td>
<td>-------------------------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>11</td>
<td>Consistent terminology: Having the same name throughout the systems for all medical concepts.</td>
<td>See design principle 2.</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Format text to visually associate drug categories: For example capitalizing the first letter for drug brands and lowercase for substances.</td>
<td>There is a clear differentiation between drug brands and substances. Drug brands have the first letter capitalized and lowercase for substances.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Emphasize differences in similar drug names: Avoiding having look-or-sound-alike drug names adjacent in lists.</td>
<td>There are no differentiations between similar drug names.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Clearly legible font: Use sans serif fonts, size 10 or 11 preferably.</td>
<td>Sans serif fonts have been used and the font size is most often size 10 or larger in order to emphasize headings.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Unambiguous units: Placement of units closely adjacent to values.</td>
<td>Units are unambiguous.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Manageable pick lists: Break long lists into sections.</td>
<td>Not applicable since no extensive pick lists are available. The only list appearing is when searching for a substance or drug brand name (see figure 10). However the list is not too long and does not need a breakdown into sections.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Multiple entry options: Limit time by start and stop times or duration.</td>
<td>Not applicable to Sfinx and NjuRen.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Concise language: Place significant words first and details later. Display ten words or less in a recommendation and provide a link to the full text.</td>
<td>The recommendation and clinical consequence is concisely put and the user is able to ‘Read more’ if necessary. The recommendations are sometimes longer than 10 words but still short and concise. Most of the participants agreed upon that the language was concise and the most</td>
<td>P</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Results from evaluation using general usability heuristics (18) (P – think-aloud participants, D – developers).
<table>
<thead>
<tr>
<th>No</th>
<th>User interface design principle</th>
<th>Results from validation</th>
<th>Method</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td><strong>Representational formats:</strong> Show trends in graphs rather than tables.</td>
<td>Essential information was presented first. However, one participant questioned the general warnings appearing at the top of both systems. The participant argued that these should be shortened to reduce the space it uses. The participant did nonetheless question if this was possible since the warnings are important.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td><strong>Visually distinct screens for confusable items:</strong> Layout or colors should be different for confusable items.</td>
<td>All of the participants agreed upon that there are no confusable items. One participant did however emphasize the importance of not taking number for facts in NjuRen. The same participant did also say that it was less confusing and easier to understand the calculated values in NjuRen thanks to the small graph showing three different values based on diverse calculation methods.</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td><strong>Custom order sets:</strong> The possibility to customize by the clinician or institution.</td>
<td>The only thing that is customizable is which of the functionalities in the Janus toolbar is to be shown or not. For instance it would not be useful to have access to the elderly-drug checking in a maternity ward.</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td><strong>Clinical context:</strong> Show relevant patient-specific information.</td>
<td>Sfinx accessed through the EHR takes into account which medications are present in the EHR, thus creating a list of patient-specific DDIs. NjuRen also takes into account medications and other personal information which is clearly stated in the system. One participant expressed that NjuRen should extract more parameters from the EHR than it does today, especially length and weight of the patient. The participant said that it was nice to have more patient-specific information as it appeared in this new prototype.</td>
<td>P, D</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td><strong>Active order forms:</strong> Content, layout, instructions can be automatically modified by patient-specific EHR data.</td>
<td>Not applicable to Sfinx and NjuRen since they are not embedded into order forms.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td><strong>Log analysis:</strong> Keeping logs in order to periodically analyze for instance consistent alert overrides.</td>
<td>The only logs that are kept are how many times the users at each clinic click the Sfinx or NjuRen button in the Janus toolbar and of which medications users have clicked and read more.</td>
<td>D</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Results from evaluation using design principles for medication-related CDSSs (9) (P – think-aloud participants, D – developers).
<table>
<thead>
<tr>
<th>No</th>
<th>User interface design principle</th>
<th>Results from validation</th>
<th>Method</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>Tiered severity level:</td>
<td>Not applicable to Sfinx and NjuRen since they do not have interruptive alerts.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interruptive alerts should be reserved for the two or three highest levels of severity to minimize impact on workflow.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Concise text, justification:</td>
<td>Both Sfinx and NjuRen have a concise text at a first glance of each interaction or warning. Sfinx has it in its detailed list where you can choose to read more via link. NjuRen provides a brief recommendation under 10 words in the first list presented with a link to the complete recommendation further down.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Content of the alert should be limited to 1-2 lines with a justification separated by white space. Supporting evidence may not be necessary to view on the primary alert interface as long as information is easily accessible via links.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Clear response buttons:</td>
<td>Not applicable to Sfinx and NjuRen since they have no order functionalities.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>There should be buttons for order or cancel available with simple labels and action links to additional options. The default action, achieved by for instance pressing the Return key, should be the preferred action.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Concurrent alert priority:</td>
<td>A clear priority structure has been applied in both Sfinx and NjuRen. The most serious ones are marked in red for instance and always presented first. These are also the ones determining which color the button is going to turn into in the Janus toolbar. The low-severity alerts, colored green, comes at the very bottom of the list of interactions and alerts. Several participants said that this prioritization was effective and significant.</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple alerts should be prioritized, emphasizing high-severity and deemphasizing low-severity alerts.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Unobtrusive reminders:</td>
<td>Not applicable since Sfinx and NjuRen does not include reminders to take certain actions.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>May be designed as flags in names lists; prioritized and color-coded messages in reserved screen areas.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Meaningful color sets:</td>
<td>The participants felt at home with the color set. The color set utilized was similar to the design principle but instead of orange there was yellow and instead of yellow the system uses white. The color set is easy to understand and consistent throughout both systems which is significant since both of them belong to the Janus toolbar which should have consistency amongst each decision support tools.</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum 5-6 colors matched across all systems. Use color shades for gradient. Red (high priority), orange (medium priority), yellow (low priority), green (safe conditions), blue (informational messages), grey (unavailable).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td><strong>Text luminosity:</strong> Dark text on light background, high contrast ratio and match appropriate color pairs.</td>
<td>Text luminosity is good in both systems.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td><strong>Filtering rules:</strong> Increase specificity by evaluation of more EHR data in trigger rules and suppress “false positives”.</td>
<td>This design principle is about interruptive alerts and thus not applicable to Sfinx and NjuRen.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td><strong>Curate, revise trigger rules:</strong> Periodic reviews of frequently overridden alerts.</td>
<td>Not applicable since alerts are not overridden, they are simply ignored.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td><strong>A DDI alert should include the following information:</strong> 1) Drugs involved, 2) Seriousness, 3) Clinical consequences, 4) Mechanism of interaction, 5) Contextual information, 6) Recommended action(s) and, 7) Evidence.</td>
<td>As seen in figure 11, the DDI alerts in Sfinx include five of the seven parts of information a DDI alert should include.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 34.1 | **Drugs involved:** It should be easy to identify which pair of drugs that interact and it would help to include both the product name as well as generic ingredients name. | Both generic and brand name is included and easy to identify. |
| 34.2 | **Seriousness:** Signal word indicating the level of seriousness. | The warning does not have a signal word but the color and the A-D categorization indicates the seriousness instead. |
| 34.3 | **Clinical consequences:** Clear description of potential adverse clinical outcomes. | The clinical consequence is clear and concisely stated. |
| 34.4 | **Mechanism of the interaction:** Description of the mechanism of the DDI. | Available via ‘Read more’ link. |
| 34.5 | **Contextual information/modifying factors:** Patient-specific factor information, such as lab results for instance, should be included in order to add relevance the alert. | No modifying factors are included. Recommendations are solely based on how different drugs interact in general. |
| 34.6 | **Recommended action(s):** Providing guidance on actions that could lessen the potential harm. | Recommended actions are clearly stated. |
| 34.7 | **Evidence:** Evidence for the DDI alert should be presented with information | The strength of the source of it is clearly stated by the numbering of the A-D categorization. However further evidence |
about the strength and source of it. for the alert is only accessible by clicking the ‘Read more’ link.

35 **Consistent terminology**: Consistency throughout DDI decision support systems. Terminology is consistent throughout the system.

36 **Timely presentation of DDI alert**: Display alerts at the point of decision making. The colors of the Janus toolbar activates when the user has opened the patient’s medication list in the EHR. Also when the user creates a new prescription the color and information on the buttons will update and change depending on the new prescription. Thus it activates in time for any action regarding medications. The timeliness could not be tested in the think-aloud since the prototypes could not be viewed going through the actual EHR.

37 **Resolving DDI alerts**: Resolving a DDI alert should be possible through performing as few steps as possible. There could also be a list of actionable choices to perform directly and continue. Not applicable to Sfinx and NjuRen since there are no interruptive alerts to resolve.

38 **Override reasons**: The possibility to reject the alert and provide rationale for rejection should be available. Not applicable to Sfinx and NjuRen since there are no interruptive alerts to override.

<table>
<thead>
<tr>
<th>No</th>
<th>User interface design principle</th>
<th>Explanation of violation</th>
<th>CDSS</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Recognition rather than recall</td>
<td>No clear button indicating where to find system documentation or help.</td>
<td>Sfinx &amp; NjuRen</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Help and documentation</td>
<td>No system documentation available.</td>
<td>NjuRen</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>Visibility of system status</td>
<td>If a commentary is sent there is a message indicating that it has been sent. However this message disappears almost immediately, giving the user no time to read it.</td>
<td>Sfinx</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Consistency and standards</td>
<td>Categorization of the warnings differs slightly and the one in Sfinx was preferred.</td>
<td>NjuRen</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3: Results from evaluation using design principles for alerts in medication-related CDSSs (9,10) (P – think-aloud participants, D – developers).

3.2.1 Summary of violations

A total of 10 direct violations against the design principles were identified and below these are presented in table 4 where they are sorted based on their severity score.
<table>
<thead>
<tr>
<th></th>
<th>Error prevention</th>
<th>The eGFR calculator allows for any value, both positive and negative thus allowing for errors to happen. By not allowing any unrealistic input errors could easily be prevented.</th>
<th>NjuRen</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Help users recognize, diagnose, and recover from errors.</td>
<td>No explanation or real error message provided when entering obvious faulty values into eGFR calculator.</td>
<td>NjuRen</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>Emphasize differences in similar drug names</td>
<td>There is no emphasis on differentiations between similar drug names.</td>
<td>Sfinx</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>Concise language</td>
<td>General warnings appeared to be long and take up much of the space. However they are significant and must stay there.</td>
<td>NjuRen</td>
<td>1</td>
</tr>
<tr>
<td>34.4</td>
<td>Mechanism of the interaction</td>
<td>Only available via ‘Read more’ link.</td>
<td>Sfinx</td>
<td>1</td>
</tr>
<tr>
<td>34.5</td>
<td>Contextual information</td>
<td>In the actual warning, no modifying factors are included.</td>
<td>Sfinx</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4: List of all direct violations against user interface design principles

3.3 Other results derived from think-aloud

The think-aloud participants were able to bring valuable input to more user interface design principles than those where their clinical expertise was absolutely necessary. This is a benefit from the think-aloud sessions. Since the researcher does not interfere it allows for the participant to freely express what they think of and focus on. This also meant that much of what was said during the think-aloud could not be related to any of the previously mentioned user interface design principles. However these commentaries were still of great value since it adds insights to things to be improved specifically for Sfinx and NjuRen meanwhile the design principles are for any medication-related CDSSs. Following are the main positive things mentioned by the participants that could not be directly related to a user interface design principle:

- It is good to have a direct link to FASS (encyclopedia for drug information in Sweden).
- Sfinx and NjuRen save a lot of time since they provide clear and concise information regarding the drug-drug interactions and prescription support for renal insufficiency. The alternative is to look up each drug in FASS.
- It is good to have the ability to choose different eGFR intervals. This is useful if a patient is on the edge of an interval or in between.
- It is good that you can expand the graph in Sfinx and see both brand names and substances. However this increases the size of the graph even further and makes it even more difficult to get an overview.
• It is easy to use both Sfínx and NjuRen when you are able to first get an overview. Then you can click on an interaction or to ‘Read full recommendation’ and be sent down to it in the system where it is also marked.
• All the information you need is provided in Sfínx.
• It is nice to have separate windows for system documentation and further readings.
• Most parts of Sfínx and NjuRen are really nicely designed and easy to understand.

In general the participants where quite positive to the outlook of both Sfínx and NjuRen and they solved the simple tasks given them easily. However they had some concerns and those were:

• The graph in Sfínx could be troublesome and not as useful if the patient has a very large list of medications since then all of them could not fit into one screen. This means that you lose the main benefit of it, which is a quick overview.
• When entering NjuRen you would want to know the renal function and the eGFR first. However currently the calculated eGFR is quite small and not as prominent as wanted.
• When you send a commentary it is not stated who receives it.

The issue with the graph is a cosmetic issue and difficult to deal with because not matter how one designs the user interface it cannot fit an infinite amount of DDIs altogether on the screen at once. If the medication list is long the graph simply loses some of its benefits. However those DDIs that do appear without the need of scrolling are the most severe and participants expressed that most often they do not even look at the interactions classified as white (B) or green (A). Thus it becomes a cosmetic issue which may need further development in the future but it is not necessary for now.

The problem with who the recipient is of the commentaries is also a minor cosmetic issue. It is stated in both NjuRen and Sfínx who the responsible developer and owner of the systems is. Thus it could be assumed that they receive the commentary. However this information is stated at the very bottom of the systems and it could be made more prominent. It would be a simple change to add who the recipient is in the window where a commentary is written.

The issue of emphasizing the eGFR in NjuRen is a minor problem since the eGFR is clearly stated. On the other hand such significant information should not have to be looked for. As stated in user interface design principle 18 in table 2, significant words should be placed first, so should the eGFR since it is the core value of which recommendations are based upon. Currently the general warning takes up most of the space and focus is put on them and values included from the EHR. These are significant and this has been recognized by the participants and thus it becomes a trade-off, whether to keep the warnings prominent since they should not be missed or to de-emphasize them and make the eGFR value more prominent.
4. Discussion

4.1 Major findings

This study did successfully achieve its aim of validating Sfinx and NjuRen using user interface design principles as a frame of reference. The aim was ultimately fulfilled by answering the main research question which was formulated as such: ‘How well does the user interface design of medication-related CDSSs Sfinx and NjuRen live up to both generic and specific user interface design principles?’ The results indicated a total of 13 problems with Sfinx and NjuRen if one includes the additional three mentioned in chapter 3.3. Amongst these there were only two violations that were more severe, scoring a three in severity. These two were both related to help and documentation indicating that this is a main issue with these versions of Sfinx and NjuRen. The other eleven violations were thus minor and not truly urgent and necessary in order to have a well-functioning CDSS. Another outcome of the results is that Sfinx and NjuRen appeared to do better in the evaluation using more specific principles to medication-related CDSS. As seen in table 4, the majority of the violations and the most severe were all between the numbers 1-10, referring to the general design principles in table 1. One could draw the conclusion that Sfinx and NjuRen are quite well-designed in regards to their clinical content but have a few issues as a computer system in general that affects its usability.

Furthermore it appears that Sfinx and NjuRen had the same amount of issues with their user interfaces as seen in table 4. However if NjuRen had been evaluated as it appears now with buttons that does not function and so on, it would have had a much higher severity score. The outcome would have been that NjuRen should not be released yet without fixing these problems. Reflecting on this evaluation process it would perhaps have been better to wait with the evaluation of NjuRen since it was too faulty at this moment. On the other hand according to these results, the user interface design of NjuRen is not that bad if all things will function as intended. This is a good indication for the developers and their future work with NjuRen and could play a part in an iterative design process.

During the heuristic evaluation it became apparent that, as previously stated, neither Sfinx nor NjuRen had interruptive alerts. Thus the design principles from Horsky et al. (19) had to be removed. There were also some other user interface design principles in table 1 that clearly referred to interruptive alerts and this indicates that this type of alert is not uncommon in medication-related CDSSs. The question is however if it is a good thing that Sfinx and NjuRen does not have interruptive alerts or if they are missing an essential part? Interruptive alerts should be reserved for the most severe warnings but they do however still interrupt the workflow. This could cause irritation among the users and alert fatigue which leads to users closing the dialog box without reading it because it disturbs their work. On the other hand if a system has no interruptive alerts, they may go completely unnoticed. It is extremely difficult to tell what works best. However if interruptive alerts are to be included it would require high
specificity and clinical context into these alerts and this is not the case for at least NjuRen which makes it probably best to keep their warnings unobtrusive until more data can be retrieved and utilized from the EHR.

4.1.1 Recommendations for improvements

The sub-question to the main research question was ‘Are there any areas of improvements in Sfinx or NjuRen based on the user interface design principles and if so, what do the principles propose to be done?’ By reviewing all of the user interface problems with Sfinx and NjuRen, a list of recommendations for improvements was created. The improvements are based upon the recognized user interface problems, what the user interface design principles recommend and the researcher’s judgment to include unique recommendations specific for Sfinx and NjuRen. In the following list, each recommendation includes a reference in brackets to the user interface design problem it solves. If ‘other’ is put within the brackets, it refers to the other issues that came up from the think-aloud sessions presented in chapter 3.3.

1. (1) In order to keep the user informed regarding what is going on, some type of message and confirmation that a commentary has been sent should be included and it should appear in the ‘Comment on an interaction’ window. That would constitute appropriate feedback.

2. (4) The systems should be kept consistent and always mention both drug brand name and the substance, thus this should be added to NjuRen. This recommendation is also supported by user interface design principle 34.1 stating that it helps to include both. In addition to keep the systems consistent NjuRen should be revised in terms of its classification of warnings. It should be considered to add numbers to the letters as it appears in Sfinx.

3. (5) The system should prevent errors to be done by not allowing obvious faulty values in the eGFR calculator.

4. (6, 10) Actions should be visible and at the moment it is difficult for the user to find the action to read the system documentation and instructions Sfinx. This should be made visible and a link should be put somewhere in the system in order to make this action more easy to find. Additionally it is highly significant that NjuRen gets its own documentation with a clear link or button for access.

5. (9) An error message explaining what went wrong and suggesting what to change should be included in the absolute eGFR calculator in NjuRen.

6. (13) Consider using “Tall man” lettering in order to emphasize look-or-sound-alike drug names adjacent in lists. (Example: prednisone and prednisolone \(\rightarrow\) predniSONE and predniSOONE (9)).

7. (18) If there is time, consider reviewing the general warnings in NjuRen and see if it is possible to shorten these.

8. (34.4) Consider putting the mechanism of the interaction in the initial warning. However it is questionable if this is good for the design since it will increase the size of the concise warnings in Sfinx.
9. (34.5) For the future it would be good to see if more personal data could be included and used in the recommendations. However it is recognized that this change is not simple since it requires more interactions with the EHR, which is not SLL’s system, and it demands more of the recommendations which are provided by a Finnish company, not controlled by SLL either.

10. (other) Include information in the ‘Comment on an interaction’ window of who the commentary is sent to.

11. (other) Consider emphasizing the eGFR values more. This could be done by making the graph with different eGFR values larger and more prominent perhaps.

12. (other) The issue of the graph in Sfinx is quite cosmetic only something to review if there is time.

4.1.2 Previous research

The results presented in this study are quite unique to this case but other studies have been done on CDSSs with similar functionalities related to the prescribing process. For instance Smithburger et al. (31) has done a critical evaluation of a CDSS detecting DDIs. One of the main conclusions made from their evaluation was that this type of CDSS should include more patient-specific information. The clinical context was mentioned amongst the design principles for this study but this was not an issue for either Sfinx or NjuRen. Especially not currently due to the new intended changes of NjuRen where more patient-specific information is extracted from the EHR and taken into account. This study complies with the importance of clinical context but did not find the same result as Smithburger et al. (31) that this should be further improved.

Moss et al. (32) evaluated nurses’ use and perceptions of a medication-related CDSS. They found that the nurses preferred decision support that was concise, color coded and easily accessed. The color coding was something praised in this study by the participating physicians and the importance of concise language was discussed. Access was not a matter in this study since this could not truly be simulated. Apart from accessibility the results from Moss et al.’s study complied with outcomes from this study though more aspects were discussed in this study. A possible reason could be the difference of having nurses or physicians involved in the evaluation.

Lastly an evaluation of usage patterns has been done on Sfinx as previously mentioned in the introduction. This study by Andersson et al. (16) did however differ greatly from this evaluation of the user interface which makes results difficult to compare. It did however emphasize that 98% of the users that responded to their questionnaire would recommend Sfinx to their colleagues. This mainly positive attitude towards Sfinx was also evident in this study during the think-aloud sessions and interviews as noticed in chapter 3.3.
4.2 Theoretical contribution

The strength and theoretical contribution that this study brings is an example within this research field of how a medication-related CDSS could be evaluated and validated using newly developed and current user interface design principles. The result from this study highlights the fact that there exists user interface design principles specific for medication-related CDSS and how these can be applied. Thus there is no excuse to not utilize these when developing this type of system which appeared to be a problem as mentioned in the introduction. Regarding usability theory this study highlights some aspects that were not covered by the user interface design principles, especially the more specific for this type of CDSSs. For instance a discovered aspect was the importance of emphasizing the eGFR in CDSSs with dosing support for renal insufficiency.

4.3 Strengths and limitations of the study

4.3.1 Strengths

The main strength of this study is the choice of methods. Instead of relying on one or two, this study made use of several data collection techniques and methods in order to obtain a firm basis for evaluation and add to the study’s credibility. For instance a think-aloud has drawbacks in terms of the verbalizations being incomplete due to limited memory capacity of the participant that hinders the cognitive process (27). By including an interview after the think-aloud itself it was more likely to minimize this problem and obtain a fuller description of the participant’s reasoning (26).

Furthermore another strength of the study is the involvement of physicians and the addition of the think-aloud sessions as a complement in the heuristic evaluation. This enabled a more complete validation since even though a person is experienced in usability, it is complex to determine the accuracy of some design aspects without subject matter expertise. This is supported by research done by Yuan et al. (33) where they in their evaluation of a CDSS found that combining evaluators with expertise in usability with evaluators with expertise in the specific domain was most effective in a heuristic evaluation of healthcare IT. Additionally the involvement of users did add a bit of user perspective and aspects not covered by the user interface design principles which was a bonus.

Lastly a final strength is that the heuristic evaluation was performed using both general user interface design principles and those highly specific for medication-related CDSSs. On one hand a medication-related CDSS share similarities to other computer systems and CDSSs thus general design principles are always applicable. On the other hand a medication-related CDSS contains functions specific for this type of system, thus it valuable to include design principles that reflects these specific functionalities.
4.3.2 Limitations

Despite the strengths in the choice of methods and how these complement each other, there are limitations in how well they were applied. The main limitation was the issue of having only one person acting as evaluator in the heuristic evaluation. It is generally suggested and recommended to have a set of 3-5 usability experts involved in the evaluation order to find most of the problems with a user interface (23). Thus if the amount of evaluators had been increased the result section would probably had included more violations against the user interface design principles. However this does not mean that the evaluation is completely dismissible since there are several instances of heuristic evaluations where only one evaluator was involved (23). It is also significant to point out that the heuristic violations found still are problems to be fixed which is helpful in the design process of the systems. But with more evaluators involved more could have been found probably.

Moreover another limitation on this subject is that the researcher cannot be perceived as a usability expert. The researcher does however have experience in heuristic evaluation and especially the use of Nielsen’s 10 heuristics (18) from a bachelor’s degree in system science and informatics. Additionally this was another reason for involving subject matter experts in order to make up for this issue.

Further on there is always the issue of objectivity in qualitative studies since the researcher is intimately involved, thus compromising the conformability of the study. However this cannot ever be eliminated completely from qualitative research. Nonetheless it is noteworthy to state that the researcher’s identity and values has an impact on the analysis of the data but efforts have been made in order to diminish judgments and preconceptions of the topic in order to obtain some level of objectivity during the study. It is also significant to emphasize that SLL did not affect the judgment of the researcher and had little involvement overall.

The close involvement of the researcher also ties into the reproducibility of the research. If another researcher was to perform the exact same study the outcome would probably differ a bit, thus also compromising the study’s reliability. This is though another aspect of qualitative studies as it is impossible to replicate a social setting (28) such as the think-aloud sessions for instance. Thus despite of the involvement of the researcher, a qualitative study is somewhat difficult to replicate and find the similar results in general. However effort has been put into the method chapter in order to provide a fairly detailed description of the research process in order to allow for other studies to replicate this study as much as possible.

Lastly based on all the limitations of this study and the uniqueness of these two systems it is not possible to generalize from these results. Another CDSS will be designed differently and thus have other heuristic violations.
4.4 Implications of findings

The foremost implication of the study’s findings is the part it can play in the continued development of Sfinx and NjuRen. The results provide SLL with a clear list not only specified with user interface problems but also recommendations on how to overcome these. Hopefully the results aid in their work towards providing useful and effective CDSSs to the clinicians within the county.

Furthermore this study presents quite unique results since it focuses on two specific CDSSs thus compromising the generalizability. However it provides a framework of compiled user interface design principles that could be applied to any medication-related CDSS.

4.5 Suggestions for future research

If this study was to be replicated in the future it is suggested to perform the heuristic evaluation properly to overcome the main limitations of this study. This means involving 3-5 usability experts as evaluators as recommended. However it is significant to still involve subject matter experts since they bring a clinical expertise necessary for evaluating this specific type of system. Furthermore the think-aloud could be redesigned a bit or explained better. A slight problem was that the participants focused very much on the medical aspects, understandable based on their occupation, and this resulted in an abundance data not relevant to the user interface of the systems. Thus in order to save time and gain more valuable data, some redesigning should perhaps be done.

Another suggestion is to involve users that have no experience of Sfinx or NjuRen. This study was limited in time and chose to include users with experience of the old prototype in order to reduce the time spent on describing the systems and their functionality. However if a user has no experience it could add another perspective and provide insights to how well-designed and helpful the user interface design is to the novice user.

Lastly another study could be performed in the future with another focus. Despite applying the user interface design principles specific for medication-related CDSSs, it is also highly interesting to evaluate the usefulness of these guidelines. This study did not have room for this but since these guidelines are so new it could be interesting to put them to the test and see if they are truly able to identify all usability issues.
5. Conclusion

This study aimed at validating the medication-related CDSSs Sfinx and NjuRen using user interface design principles as a frame of reference. Objectives along the way was to find appropriate user interface design principles, apply them to Sfinx and NjuRen and see where they fell short and what could be done to improve the design. The outcome was that Sfinx and NjuRen had a total of ten direct violations against the 38 user interface design principles used and an additional three addressed during think-aloud sessions not related to the design principles. Amongst the violations most of them were minor or cosmetic but the two related to help and system documentation were more severe and should be dealt with soon. Furthermore the results also showed that the majority of the violations and the most severe were related to general usability heuristics indicating that Sfinx and NjuRen are quite well-designed in regards to their clinical content but lacking a bit as a computer system in general.

In conclusion, this study provided a thorough example of evaluating medication-related CDSSs that contribute to the understanding of the usefulness of applying user interface design principles to this type of system as a means to increase the system’s usability.
References


29. Terapigruppen Äldre och läkemedel. Läkemedelsgenomgångar: på praktiken i praktiken (Pharmaceutical reviews: at the practice in practice) [Internet]. Västra Götalandsregionen; 2014 [cited 2016 Apr 7]. Available from: http://epi.vgregion.se/upload/L%C3%A4kcemedel/%C3%A4ldre och l%C3%A4kemedel/Utbildning/14-09-26 L%C3%A4kemedelsgenomg%C3%A5ngar.pdf


Appendices

Appendix A – Consent form

Consent form

Study: Evaluation of clinical decision support systems NjuRen and Sfinx using user interface design recommendations and principles.

Researcher: Emma Molin, Master student at Karolinska institute and Stockholm university. If any questions: Phone: …… E-mail: ……

This study is a part of the master thesis within the field of health informatics undertaken by the researcher in question. The purpose of the study is to compare two medication-related clinical decision support systems developed by the department of E-health and Strategic IT at Stockholm county council against user interface design principles. The study aims to answer the question of how well these clinical decision support systems live up to these principles. You will contribute to the study by participating in a so called think-aloud with subsequent interview.

As a participant in this study you are completely anonymous. Only your profession will be mentioned to describe you as a participant. Anonymity is guaranteed by the researcher who will ensure that when stored, all of the collected data will solely be related to an ID-number and never any personal information about you. This protection of your privacy is has been assigned in order to protect you as an individual since the results of the study will be shared by various people at Karolinska institute, Stockholm university and Stockholm county council.

Participation is voluntary and you have the right to at any time cancel the think-aloud/interview or withdraw your consent even after signing this form.

I hereby verify that I have read the information above and give my consent to participate in the study:

______________________________________________  __________________________________________
City and date                                  City and date

______________________________________________  __________________________________________
Participant’s signature                         Researcher’s signature
Appendix B – Think-aloud

Think-aloud instructions

Your patient Anna aged 83 has arrived on her yearly check-up with a pharmaceutical review. Anna lives at a retirement home and walks short distances with a walking frame and live support, uses a wheelchair otherwise. Anna suffers from mixed dementia, asthma, angina pectoris, acid refluxes and depression. Lately she have been eating and drinking normally but at this yearly check-up she appears a bit thin. Anna responds inadequately when spoken to and her blood pressure is 130/80 with a pulse of 70. Moreover it is stated in her medical journal that she weighs 55kg and is 158cm tall.

Momentarily Anna is taking the following medications:

- Omeprazol 20mg 1x1
- Pradaxa 110mg 1x2
- Imdur 60mg 1x1
- Furix 40mg 1x1
- Isoption retard 240mg 1x1
- Risperidon 0,5mg 1x2
- Citalopram 20mg 1x1
- Ebixa 10mg 1x1
- Singulair 10mg 1x1
- Bricanyl 5mg 1x1
- Nitrolingual sos
- Impugan 20mg 1 sos
- Ventoline 0,1mg 1 sos
- Movicol 1x1-3
- Prednisolon 5mg 5x1 sos
- Acetylcysteine 200mg 1x3 sos

Your tasks:

1. The Janus toolbar have reacted to the medications prescribed for Anna and indicates that there are interactions between drugs that should be reviewed. Think now first of what information regarding the drug-interactions that you think is the most important and interesting for you. Then open Sfinx by clicking the interactions-button and try to locate this information.

2. Citalopram Sandoz and Impugan has an interaction classified as C2. Try to find out what the medical consequences and recommendations are as well as the sources for these.

3. Try to figure out what the difference is between a B1 interaction and a B0 interaction.

4. Look at the synoptic diagram of the interactions. Figure out the purpose of the tiny grey circles with numbers within them located next to the drug names.

5. Hopefully you have a good sense of Sfinx by now. Finish off with looking around one last time and see if there is anything in particular that captures your interest.
6. You are now going to look at a screenshot of how NjuRen would have looked for the same patient. This is a prototype of NjuRen and is thus flawed. Now think like you did in task 1, what information is of greatest interest to you when using NjuRen? Then see where to find this information.

7. Continue to familiarize yourself with NjuRen. See if there is anything in particular that captures your interest.

8. As a final task, you shall now open the web-browser and the NjuRen tab. This prototype is unfortunately even more flawed but it is the only functional version available at the moment. The obvious shortcomings aside, familiarize yourself with NjuRen again and take the opportunity to look around and check things that were not available in the screenshot. (This prototype does not show anything related to the given patient case).

Thank you for your participation!

Appendix C – Interview questions

1. Did you perceive anything as irrelevant in either system?
2. Was there anything that could be misinterpreted or confusing?
3. Did you ever react on the language used, did it feel appropriate?
Appendix D – Additional screenshots of Sfinx and NjuRen

Figure 7: Screenshot of NjuRen used in think-aloud sessions
Figure 8: The window in NjuRen where it is possible for the user to enter values and calculate their own absolute eGFR.

Figure 9: Classification of DDIs in the system documentation.
Figure 10: Doing a search for drug or substance in Sfnx accessed through the web

Figure 11: Example of a severe nonintrusive alert in Sfnx