Investigating Clinical Decision Support Systems
Success Factors with Usability Testing

Vitri Tundjungsari
Faculty of Information Technology
YARSI University
Jl. Letjen Suprapto, Jakarta, Indonesia

Ahmad Sabiq
Faculty of Information Technology
YARSI University
Jl. Letjen Suprapto, Jakarta, Indonesia

Abdul Salam Mudzakir Sofro
Faculty of Medicine
YARSI University
Jl. Letjen Suprapto, Jakarta, Indonesia

Aan Kardiana
Faculty of Information Technology
YARSI University
Jl. Letjen Suprapto, Jakarta, Indonesia

Abstract—Clinical Decision Support Systems (CDSS) have been used widely since 2000s to improve the healthcare quality. CDSS can be utilized to support healthcare services as a tool to diagnose, predict, as well as to provide clinical interpretation, alert, and reminder. There are many researches of CDSS implementation on literatures but not many of them present the evidence of CDSS successful implementation. In spite of the potential use of CDSS, there are some researches that reveal the failures of CDSS implementation. This paper contributes to CDSS development by investigating and exploring CDSS success factors with usability testing. The testing involves participants from different types of backgrounds (physicians, IT developers, and students). The participants are being asked to experience three different CDSS to predict cardiovascular risk factors. The result of the research shows that involving different type of users give more insight to design process. It can be concluded that user center design is very critical to produce successful CDSS.

Keywords—Clinical decision support systems; success factors; user; usability testing

I. INTRODUCTION

The background of this research starts from the widespread development and use of information technology to support decision making in the health field, also called as: Clinical Decision Support System (CDSS). This study aims to find out the understanding of CDSS from the perspectives of physicians (such as: doctors and prospective doctors/medical faculty students) as well as understanding from the information technology staffs (such as: IT developers, lecturers and students of information technology department). We also investigate the public’s understanding of clinical decision support system. The physician/medical staff is chosen as representative of the experts from the content, i.e. health perspectives, while IT developers as the representatives of expert who develop the CDSS.

This research is important to find out the extent to which CDSS usage and benefits from participants’ point of view. The focus of CDSS application in this study is for the detection of chronic diseases, i.e.: cardiovascular disease. Chronic diseases provide a greater public health burden than acute illness because it requires more visits and medications [1]. Thus the use of CDSS for chronic diseases prediction is expected to reduce the cost of treatment and finally can decrease the mortality caused by chronic diseases. In this paper we compare three CDSS to predict cardiovascular risk using usability testing. Those three CDSS have different interface and indicators to perform calculation of cardiovascular risk factors.

The rest of the paper is organized as follows. Section 2 presents a brief review of existing CDSS research. Section 3 presents the method used in this research and Section 4 describes our findings. Finally in Section 5, we conclude the paper and state our future work.

II. CLINICAL DECISION SUPPORT SYSTEM

A. CDSS Application

CDSS is a tool with electronic media used to determine diagnosis, clinical interpretation, trends, alerting, reminder, predictive analysis with applications (services or interfaces) which is connected to the data. Another definition of CDSS is system that provides information to medical personnel, patients or individuals or populations, to produce faster, more efficient, better health outcomes for both individual health services and the health of a population [2]. From the above definition, it can be concluded that CDSS has the main objective to support various clinical functions, such as: providing documentation and clinical coding, organizing clinical complexity, storing and maintaining patient databases, tracking patient orders, monitoring and tracking health condition, as well as used for preventive measures disease.

In the following discussion, we present some examples of CDSS applications that have been developed and implemented in the real world, as follows:

- ATHENA. Athena is a CDSS application developed in 2002 as a tool to provide guidelines for people with hypertension. Athena helps patients in controlling blood pressure and recommends appropriate treatment options for patients. Athena also provides information
on hypertension medications and protocols related to hypertension management. The Athena system is designed independently so that it can be integrated with various electronic medical record systems (Electronic Medical System/EMS); thus Athena can adapt in various health information systems. The effectiveness, accuracy, and success of Athena's implementation have been studied and examined in various studies [3], [4]. It can be concluded that the use of Athena supports the effectiveness of treatment for hypertension disease.

- **ISABEL.** Isabel is an application for web-based decision support system developed in 2001. Isabel can be used by patients of all ages, from birth to old age. The Isabel database provides a wide selection of major specializations, such as internal medicine, surgery, obstetrics and gynecology, pediatrics, geriatrics, oncology, toxicology and bioterrorism. Isabel produces a diagnosis for a set of clinical features, such as: symptoms, signs, test results and investigations, followed by recommended medication recommendations. The Isabel system is linked to the EMR to make it possible to extract an existing diagnosis as well as for patients with other specific data. The system also provides features to help clinicians answer their questions with the latest information and knowledge from textbooks and journals. Isabel has been extensively validated and shows good results for improving clinical cognitive abilities, improving patient safety and improving the quality of patient care [5].

- **LISA.** Lisa is a CDSS consisting of two main components: (1) A decentralized Oracle database: it contains all patient information about the treatment schedule, blood test results and toxicity, prescribed dosage of medication. The Lisa database is accessible to health professionals from different sectors and locations; (2) Web-based decision support module, using PROFA technology as an application development guide. In this module contains information on treatment dose advice and focused on long-term care related to disease whose treatment doses should be monitored and adjusted continuously; as the effects of treatment vary between patients with each other (e.g. for chemotherapy treatment), Lisa is important for different type of therapy, as many cases of dosage errors in medical practice [6]. Author is [7] mention that Lisa has been evaluated and accepted by clinicians.

### B. Relevant Research: Assessment of CDSS Application

Discussion of some of the above CDSS applications provides an understanding that computer-based CDSS has been applied since the early 2000s, even since the 1970s computer-based CDSS has been developed extensively [8]. Coeira in [8] also discusses several categories of computer-based CDSS based on the following objectives:

- Increase patient safety. CDSS has benefit for healthcare services by reducing medical errors, avoiding medical advice as opposed to protocols, improving treatment sequences and tests.
- Improve the quality of health services. The usage of CDSS provide benefit by increasing the service time directly to patients, improving the regularity of use of guidelines and clinical procedures, accelerating and improving the use of the latest clinical findings, improving clinical documentation and patient satisfaction.
- Improve the efficiency of health services by reducing service costs by processing faster demand, reducing examination costs due to duplicate tests, reducing medical advice errors, repetition or error treatment patterns, and promoting the use of low-cost generic drugs that have similar effectiveness with non-generic drugs.

Fig. 1 shows the results of research on the benefits of CDSS conducted by the State of Victoria in Australia (Victorian State) [2]. It shows that top three benefits of CDSS from are: (1) reduction of time in health-care processing, (2) increasing the order of treatment related to administrative issues, (3) lowering the cost of paper forms. This shows that CDSS has a lot of potential for clinicians and health care providers to make decisions, although there are still many doubts from clinicians to use CDSS.

Despite of many potentials generated by CDSS as we have previously discussed, the CDSS impact on daily clinical practice is less likely to produce the desired results [8], [9]. Therefore building a CDSS is a challenging process because theoretical knowledge is not sufficient in its implementation in the field. Thus, this study is conducted to investigate the users' understanding of CDSS’s benefits, but also to identify what factors lead to CDSS success and failure. Furthermore, the results of the study are expected to improve the efficiency of CDSS development and use in the practice of daily health services.

Other research also mention that the CDSS has been developed since 40 years ago, but the use of CDSS is still not fully utilized because it still does not meet the expectations of the organization [2]. Bright et al also mention that despite the argument that the use of CDSS can improve the quality of health services but evidence supporting this is lacking [9].

There is a research related to CDSS implementation that has been published in the Victorian Health Design Forum Report [2] discusses some of the weaknesses of CDSS in the area of clinical decision-making. The biggest factor that makes clinicians often hesitates or reluctant to use CDSS because clinicians trust their decision on “computer-based technology,” but clinicians are still the ones responsible for clinical decisions, based on available information. Another CDSS weakness is the high price of CDSS systems, including development costs, maintenance costs, and training and support costs. Time required by the CDSS to interpret the data also contributes to CDSS weakness, and also the system can be accessed by multiple login so that the system must be developed into a trusted and effective system.
A Dutch CDSS study found that 65 percent of respondents (doctors and non-doctors) say that CDSS can make mistakes; 90 percent of respondents said that suggestions and recommendations produced by the CDSS should be checked again; and 79 percent said that those responsible for patient care were clinicians, not CDSS [2]. Nevertheless, 80 percent of respondents say that CDSS can generate suggestions and recommendations that help the performance of clinicians. Fig. 1 shows the benefits of CDSS based on research result.

Another study conducted by Bright et al. found that both commercial and local CDSSs were effective in improving the health care processes, however evidence of the advantages of CDSS use in terms of clinical, economic, workload and work efficiency decisions varied greatly [9]. Thus the impact of CDSS should still be examined and evaluated further.

Author in [10] found that about 45 percent of computer-based medical information systems fail because of the user’s refusal to use them, even though technology in the system has been developed comprehensively. Other causes of CDSS failure include lack of computer skills, lack of motivation to change or add CDSS features, and loss of professional autonomy to maintain the use of CDSS. There is also evidence that CDSS services are not always used while available, due to lack of motivation from physicians to use them in clinical decision making [11].

There are related evidences that the usability of CDSS should be evaluated since it is determined as the successful implementation of CDSS [12], [13]. Authors in [12] evaluate usability of CDSS by conducting two phases of usability testing. Phase one employed think-aloud method to investigate positive outlook of the CDSS; while phase two evaluate the improvements of the CDSS based on the phase one feedback. The result of the study shows that usability evaluation using think-aloud protocol analysis and near-live clinical simulation are very successful as assessment method in order to refine the CDSS’s usability and workflow. Authors in [13] find out that CDSS evaluation using usability engineering principles is very important to identify interface problems. The usability evaluation conducted by Graham et al. [13] involves emergency physicians which then analyzed with structured method. They conclude that, user involvement should be putting into the early stage of software design.

Based on several findings of CDSS benefits and weaknesses, we carried out this research to investigate the success and failure factors of CDSS application, as discussed further in Sections 3 and 4.

### III. Method

The research’s design is descriptive analytical resulted from a qualitative research. The data obtained from 45 participants involved in usability testing using questionnaire and three CDSS application as instruments. To find the potential problems and benefits in CDSS applications, we employ usability testing. Usability defined as to what extent a product can be used by specified users to effectively, efficiently, and satisfactory perform tasks in order achieving specified goals [14].

After the usability testing performed, we gather 10 respondents in a Focus Group Discussion (FGD). In general, the process in this research is divided into three stages, i.e.: (1) data collection, (2) data processing, and (3) analysis and presentation of data.

#### A. Data Collection

The data collection method in this research uses survey methods with questionnaire as instruments. This is done because the survey results can provide data such as: behavior, feeling, trust, behavior, knowledge, ownership, personal characteristics, as well as other descriptive matters. The survey results may also provide data on the association. Questionnaires were distributed directly to respondents by collecting selected respondents and asking the respondents to use CDSS applications and fill out questionnaires based on behavioral scale. The Likert scale is used to find out the respondent’s behavior using ordinal with value 1 to 5, i.e.: 5 for strongly agree, 4 for agree, 3 for hesitate, 2 for disagree, and 1 for strongly disagree.

Focus Group Discussion (FGD) is conducted after the survey method is completed, participants invited in the FGD are those included in the sample. FGD is a research technique for collecting data through individual comments and group interactions for a specific topic. FGDs are used to test the consistency of answers in the previous survey. Thus, the FGD is useful for data collection and simultaneously used as a validation of the questionnaire results, moreover FGD is useful to explore the thoughts of group of individual. Authors in [15] suggest that a sufficient number of efficient FGDs are three. However, in this research we involve 10 participants in our FGD.

The participants in this research consist of eight physicians, five IT application developers, and thirty two students from various background (we consider these students represent public). Thus, in total there are forty five participants required to take part in the study. Participants would also be required to provide demographical information, such as: age, gender, occupation, number of years in working...
experience, highest educational degree obtained, information and computer literacy, and knowledge of CDSS. An initial meeting with the participants to introduce the research objective was held before the research. After that, participants are given access to experience three CDSS application for cardiovascular risk prediction. The discussion result of the research is presented in Section 5 of this paper.

B. Data Processing

After the data has been collected then it processed using statistical software. The qualitative data obtained from FGD results. Although FGD also has a biased tendency (e.g. one’s opinion may be influenced by group opinion or vice versa), we still choose this method because we are interested in individual opinions and group opinions through interaction among respondents. In a group, one’s opinion can trigger thoughts and ideas from other individuals so that they can dig in more input than if only interviewed individually. The questions in FGD only focus on the experience of the participants (usability testing) in using CDSS, and explore participants’ opinions regarding the advantages and disadvantages of CDSS.

The question in the survey also should fulfill reliable measurement scale with Alpha-Cronbach value > 0.70 [16]. A measurement instrument is said to be reliable when it gives a consistent score result on each measurement. A measurement may be reliable but invalid, but an unbiased measurement is not valid if it is not reliable. This means that reliability is a necessary condition but not sufficient for validity. With reliability analysis, it can be seen the relationship between the question items in the questionnaire. Alpha-Cronbach value is used as internal consistency index of the overall measurement scale, thus the items in questionnaires which less then Alpha-Cronbach Value should be revised or deleted.

C. Data Analysis and Presentation

The data gathered from the research then analysed using Wilcoxon signed-rank test. Wilcoxon signed-rank test is non-parametric statistical hypothesis test used for comparing two related samples. By using Wilcoxon signed-rank test, we can find out the preference of participants. In overall, the questionnaire consists of five parts, i.e.:

- **Part 1** participant’s profile: name, occupation/job description (e.g. physician, IT developer, student), age, gender, last education obtained, years of experience in last occupation.

- **Part 2** information and computer literacy: skills in using computer, time in using the Internet, knowledge of CDSS.

- **Part 3** experience of users while using CDSS application during usability testing: how the CDSS application support the work, the benefits and the weaknesses of the CDSS application to support the daily work.

- **Part 4** user’s feedback of CDSS application, for components: readability, content design, navigation and help, efficiency and flexibility, and error recovery.

- **Part 5** user’s opinion of CDSS application after using it

IV. RESULT AND DISCUSSION

We recruit the participants of our research from different occupation and education background, i.e.: physicians, IT developer, also students from medical and IT school from our private university, as shown as in Fig. 2. The participants in our research consist of 45 people, with gender composition of 78% female participants and 22% male participants, as described in Fig. 3.

The age of range of participants are varied from 19 to 35, with mean 22.03 years old (deviation standard 3.745). All of participant has no background knowledge of CDSS, although all of them have good computer and internet literacy. In a week, the participant using internet averagely 26.45 hours, minimum 2 hours and maximum 144 hours per week.

After the participants fill in the demographical information, then they are being asked to experience the CDSS applications. We provide three CDSS application, namely: Application A, Application B, and Application C. Fig. 4, 5, and 6 shows the applications’ interface.
A. Application A

This application is developed to calculate cardiovascular detection risk for individual with range of age 30-75 years old, and never been diagnosed atherosclerotic disease. This application has result in percentage as a result of calculation from Framingham, Joint British Societies (JBS)/British National Formulary (BNF), and ASSIGN.

This application uses six variables to calculate cardiovascular risk factors, i.e.: (1) age, (2) gender, (3) smoking status, (4) systolic blood pressure (mmHg), (5) total cholesterol (mmol/L), (6) HDL (mmol/L). The prediction is a result of calculation of ten years period from present time. Fig. 4 shows the interface of application.

B. Application B

Application B is a calculator to predict cardiovascular risk based on guidance from The National Heart, Lung, and Blood Institute (NHLBI) in the United States. NHLBI develops research, training, and education program to increase public health awareness and public health quality, related to heart, lung, and blood disease all over the world.

This application uses seven variables, i.e.: (1) age, (2) gender, (3) total cholesterol, (4) HDL, (5) smoking status, (6) systolic, (7) hypertension treatment status. The prediction is a result of calculation of ten years period from present time. Fig. 5 shows the interface of application.

C. Application C

Application C is kind of CDSS to predict cardiovascular risk based on guidance developed by The National Vascular Disease Prevention Alliance (NVDPA) Australia. NVDPA is an organization as a result of joint collaboration from four important health organizations, i.e. Diabetes Australia, the National Heart Foundation of Australia, Kidney Health Australia, and the National Stroke Foundation.

This application uses eight indicator variables to calculate risk factor of cardiovascular disease, which are: (1) age, (2) gender, (3) systolic blood pressure, (4) total cholesterol, (5) HDL cholesterol, (6) smoking status, (7) diabetes status, (8) ECG LVH. Fig. 6 shows the interface of application.
Table 1 present the variable indicators used by each CDSS application and top five opinions from the participants gathered from open question in the questionnaire and validated through FGD. The table compares three CDSS application, based on the variables and opinion gathered from participants.

<table>
<thead>
<tr>
<th>App</th>
<th>Variables</th>
<th>Opinion from participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1. Age</td>
<td>1. The user interrupts with the advertisements</td>
</tr>
<tr>
<td></td>
<td>2. Gender</td>
<td>2. Not enough explanation to use the application</td>
</tr>
<tr>
<td></td>
<td>3. Smoking status</td>
<td>3. The design of application should be more attractive</td>
</tr>
<tr>
<td></td>
<td>4. Systolic blood pressure (mmHg)</td>
<td>4. The result of calculation is not easy to read</td>
</tr>
<tr>
<td></td>
<td>5. Total cholesterol (mmol/L)</td>
<td>5. Too many inputs (more complex than others)</td>
</tr>
<tr>
<td></td>
<td>6. HDL (mmol/L)</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1. Age</td>
<td>1. Enough explanation to use</td>
</tr>
<tr>
<td></td>
<td>2. Gender</td>
<td>2. The application has feedback as expected</td>
</tr>
<tr>
<td></td>
<td>3. Total cholesterol</td>
<td>3. The application provide sufficient details of inputs</td>
</tr>
<tr>
<td></td>
<td>4. HDL</td>
<td>4. The design of application should be more attractive</td>
</tr>
<tr>
<td></td>
<td>5. Smoking status</td>
<td>5. The application is quite simple and easy to use</td>
</tr>
<tr>
<td></td>
<td>6. Systolic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7. Under treatment of hypertension status</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1. Age</td>
<td>1. Best design among others</td>
</tr>
<tr>
<td></td>
<td>2. Gender</td>
<td>2. Simple and attractive design</td>
</tr>
<tr>
<td></td>
<td>3. Systolic blood pressure (mmHg)</td>
<td>3. The units of measurement should be compatible with general use</td>
</tr>
<tr>
<td></td>
<td>4. Total cholesterol</td>
<td>4. The accuracy of the result should be tested</td>
</tr>
<tr>
<td></td>
<td>5. HDL cholesterol</td>
<td>5. Support physician work</td>
</tr>
<tr>
<td></td>
<td>6. Smoking status</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7. Diabetes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8. ECG LVH</td>
<td></td>
</tr>
</tbody>
</table>

We also rank the CDSS application based on the questionnaire, using Wilcoxon signed-rank test. We use Wilcoxon signed-rank test to compare the rank between two application, i.e.

1) Application B and C
2) Application A and B
3) Application A and C

Wilcoxon signed-rank test result between application B and C:

\[ H_0 : \eta_{\text{cvdrisk} \, B} = \eta_{\text{cvdcheck} \, C} \]

Has given value of \( z = -1.441 \). Thus, we do one-tailed test, having hypotheses as follows:

\[ H_1 : \eta_{\text{cvdcheck} \, B} \neq \eta_{\text{cvdcheck} \, C} \]

We find that \( p-value \) equal to \( 0.150/2 = 0.075 \), which is not less than \( \alpha=0.05 \), thus \( H_0 \) is accepted.

Conclusion: \( \eta_{\text{cvdrisk} \, B} = \eta_{\text{cvdcheck} \, C} \) or the median score from participants for application B is equal to application C.

Wilcoxon signed-rank test result between application A and B:

\[ H_0 : \eta_{\text{cvdrisk} \, A} = \eta_{\text{cvriskcalculator} \, B} \]

Has given value of \( z = 0.961 \). Thus, we do one-tailed test, having hypotheses as follow:

\[ H_1 : \eta_{\text{cvriskcalculator} \, A} > \eta_{\text{cvdcheck} \, C} \]

Thus, \( p-value = \frac{0.337}{2} = 0.1685 \), which is not less than \( \alpha=0.05 \), thus \( H_0 \) is accepted.

Conclusion: \( \eta_{\text{cvdcheck} \, C} = \eta_{\text{cvriskcalculator} \, A} \)

Has given value of \( z = 2.308 \).

Thus, we do one-tailed test, having hypotheses as follow:

\[ H_1 : \eta_{\text{cvdcheck} \, C} > \eta_{\text{cvriskcalculator} \, A} \]

Thus, \( p-value = \frac{0.021}{2} = 0.0105 \), which is less than \( \alpha=0.05 \), therefore \( H_0 \) is rejected.

Conclusion: \( \eta_{\text{cvdcheck} \, C} > \eta_{\text{cvriskcalculator} \, A} \) or the median score from participants for application C is bigger than application A.

From all of the test result above, we then validated the result by performed FGD. The FGD consists of 10 participants which randomly chosen from previous participants. It is confirmed that CDSS C is the best CDSS application chosen by the participants, followed by application B, and A. Fig. 7 shows the graphical information of CDSS usability comparison result.

![CDSS Usability Comparison](image_url)

Fig. 7. Participants preference of three different CDSS.
V. CONCLUSION

This research provides contribution to evidence-based usability knowledge for CDSS. As the result of this work, we can develop usability guidelines and knowledge in order to implement a successful CDSS. We also identify several components that should be carefully considered and designed to optimize the benefit of CDSS.

The most important thing in designing CDSS is to understand who the users are. In this research, we have three different types of users, i.e.: physician, IT developers, and students as representative of public. Those three different types of users have different focus in assessing CDSS’s usability, as shown in Table 2.

<table>
<thead>
<tr>
<th>Physicians</th>
<th>IT developers</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>The accuracy of CDSS result</td>
<td>The design of CDSS</td>
<td>Sufficient explanation to use the CDSS</td>
</tr>
<tr>
<td>Support physician’s work in diagnosing patients</td>
<td>Simplicity and easy to use</td>
<td>Provide feedback while experiencing the CDSS</td>
</tr>
<tr>
<td>Unit system should be compatible with environment</td>
<td>The design should be attractive to use</td>
<td>Less complexity (less input), user friendly</td>
</tr>
</tbody>
</table>

Based on the findings in our research, we can conclude that the most important thing to implement a successful CDSS knows the users. Each type of user has different focus while experiencing the CDSS. Therefore, physicians, CDSS developers, patients, and public should be included in user center design in iterative design processes to share opinion and recommendations.

Our future work is to develop the CDSS application using guidelines that has been found from this research. The application then will be tested in real environment as comparison to this research.

ACKNOWLEDGMENT

This work was funded by the Ministry of Research Technology and Higher Education Republic of Indonesia, under PUPT research schema. We would like to thank also to Universitas YARSI for its support during the research study.

REFERENCES