Assessing Acceptance Level of a Hybrid Clinical Decision Support Systems

Georgy KOPANITSA[a,1], Ilia V. DEREVITSKII[a], Daria A. SAVITSKAYA[b] and
Sergey V. KOVALCHUK[a,b]

[a] *ITMO University, 49 Kronverskiy prospect, 197101, Saint Petersburg, Russia*
[b] *Almazov National Medical Research Centre, 2 Akkuratova st., 197341,
Saint Petersburg, Russia*

Abstract. We present a user acceptance study of a clinical decision support system (CDSS) for Type 2 Diabetes Mellitus (T2DM) risk prediction. We focus on how a combination of data-driven and rule-based models influence the efficiency and acceptance by doctors. To evaluate the perceived usefulness, we randomly generated CDSS output in three different settings: Data-driven (DD) model output; DD model with a presence of known risk scale (FINDRISK); DD model with presence of risk scale and explanation of DD model. For each case, a physician was asked to answer 3 questions: if a doctor agrees with the result, if a doctor understands it, if the result is useful for the practice. We employed a Lankton's model to evaluate the user acceptance of the clinical decision support system. Our analysis has proved that without the presence of scales, a physician trust CDSS blindly. From the answers, we can conclude that interpretability plays an important role in accepting a CDSS.

Keywords. CDSS, user acceptance, data-driven, rule-based

1. Introduction

Clinical decision support systems (CDSS) are made to support evidence-based patient care and shared-decision making to improve health and wellbeing of patients. While various studies have demonstrated that CDSSs decrease medical errors and improve clinical outcomes, we can see that CDSSs did not yet reach their full potential due to low acceptance and adoption [1,2]. Among the factors that influence adoption and acceptance we can name relevance of the provided information and the validity of the system [3]. Validity and ability to interpret the decision support output can be especially problematic for data-driven CDSSs [4]. One of the approaches to solve the interpretability problem is a hybrid approach [5] where a data-driven decision support is complimented by rule-based methods and scales. We have implemented a CDSS for diabetes complications management using a three-stage hybrid approach [6]. The goal of this study is to understand how a combination of data-driven decision support methods with a rule-based interpretation affect the acceptance and adoption by doctors.

---

[1] Corresponding author, Georgy Kopanitsa, ITMO University, Saint-Petersburg, Russia, E-mail: georgy.kopanitsa@gmail.com

2. Method

2.1. Decision support system

The CDSS in focus of this study predicts 5 years risks of type 2 diabetes (T2DM) mellitus complications [7]. It includes machine learning based inference along with a FINDRISK scale [8]. The model that is the basis of the CDSS does not require sophisticated medical tests and provides the following prediction efficiency: sensitivity of 76.0% and specificity of 60.2%. The interface of the CDSS with a synthetic data is shown in the figure 1.

Figure 1. CDSS interface

The structure of the study is based on the theory of planned behavior (TPB). It considers attitude, subjective standards, and perceived behavioral control influencing behavioral intentions (and actual behavior).

2.2. CDSS efficiency

We estimated a perceived usefulness of the systems. This metric represents a degree, to which users suppose that utilizing a decision support system will increase their efficiency. We have conducted a survey with physicians who have experience of operating the system. The survey was structured into two phases. The evaluate the perceived usefulness we randomly generated CDSS output in three different settings:

- (A) Data-driven model output,
- (B) Data-driven model with a presence of FINDRISK scale
- (C) FINDRISK scale and explanation.

For each type of settings, we randomly generated a questionnaire with synthetic T2DM cases and questions. Each case was presented to a physician with a patient's basic information and vital signs including antithrombotic therapy (AH), physical activity, blood sugar, short hereditary anamnesis, blood pressure) and a setting specific (A-C) CDSS output.

For each case, a physician was asked to answer 3 questions:

- if a doctor agrees with the result,
- if a doctor understands it,
- if the result is useful for the practice.

All the questions could be answered using a Likert scale with 5 points from 1 (strongly disagree) to 5 (strongly agree).

2.3. User acceptance

The acceptance of the decision support system was evaluated using a Wilson's model of electronic health solutions' acceptance modified by Lankton [9]. The model enables assessing the following metrics: behavioral intention to use (BI), intrinsic motivation (IM), perceived ease-of-use (PEOU), and perceived usefulness (PU) of the decision support system. We measured BI and PU using 2 objects for each metric. IM and PEOU metrics were measured with 3 objects. To rate each item, we applied Likert scale with 5 points: from 1 (strongly disagree) to 5 (strongly agree):

1. Behavioral intention to use
   a. I will use the CDSS to have a second opinion on the patient's risks
   b. I believe I will utilize the CDSS in my practice
2. Intrinsic motivation
   a. The CDSS helps me to make better informed decisions
   b. I trust the CDSS as it provides interpretations of the output
   c. I trust the system as it provides references to the standard scales
3. Perceived ease of use
   a. The CDSS outcomes are clear and understandable
   b. The interpretations are clear, and I understand the reasoning
   c. The visualizations are well-defined, and I don't spend much time on their interpretation
4. Perceived usefulness
   a. CDSS improves the effectiveness of managing risks of patients
   b. It explains me why a certain risk assessment is done

After we collected and analyzed the results of the user acceptance evaluation, we have organized a study to deeper understand the reasoning of the doctors when working with the CDSS. We designed a study as a series of semi-structured one-to-one interviews with an interview script [10], which was created and approved by the research team.

1. Can you understand a model output without interpretations?
2. Are you convinced with the interpretations that the system provides?
3. Do you require an interpretation to critically assess a model output?
4. Does a reference to a scale facilitate assessment of a recommendation?
5. Can you please give any improvement comments or suggestions?

3. Results

We have gathered 161 answers with equal distribution for each setting: 53, 55, 53 for A, B, C, respectively (Table 1).

Table 1. Case scoring mean (95% confidence interval)

	Agree	Understand	Use
Setting A	4.05 (3.78, 4.32)	4.64 (4.42, 4.85)	3.8 (3.05, 4.80)
Setting B	3.16 (2.94, 3.38)	3.98 (3.66, 4.29)	3.38 (3.11, 3.64)
Setting C	3.41 (3.15, 3.67)	4.24 (3.90, 4.54)	3.52 (3.20, 3.85)
All settings	3.54 (3.38, 3.69)	4.28 (4.12, 4.45)	3.56 (3.39, 3.73)

The median values for behavioral intention to use, intrinsic motivation, perceived ease-of-use, and perceived usefulness (PU) demonstrated a general acceptance of the CDSS by the users (Table 2).

Table 2. CDSS acceptance metrics

Metric, Item	Median	Max	Min
1. Behavioral intention to use	3	5	2
1a. I will use the CDSS to have a second opinion on the patient's risks	3	5	3
1b. I believe I will utilize the CDSS in my practice	3	4	2
2. Intrinsic motivation	3	4	2
2a. The CDSS helps me to make better informed decisions	3	4	2
2b. I trust the system as it provides interpretations of the results	3	4	3
2c. I trust the system because it provides references to the standard scales	3	4	2
3. Perceived ease of use	4	5	2
3a. The model outcomes are clear and understandable	4	5	3
3b. The interpretations are clear, and I understand the reasoning	4	4	2
3c. The visualizations are clear and I don't spend much time on their interpretation	4	5	3
4. Perceived usefulness	4	5	2
4a. CDSS improves the effectiveness of managing risks of patients	4	4	2
4b. It explains me why the a certain risk assessment is done	4	5	3

4. Discussion and Conclusions

The analysis showed the highest scores were obtained in Setting A, while the lowest is obtained in Setting B. Our interview analysis has proved that without the presence of a standard scale, a physician trust blindly a CDSS results. This can increase type I errors, which can be lowered in comparison to the basic scales. We have analyzed the answers of the participating doctors to understand the reasoning behind the acceptance evaluation. From the answers of the doctors, we can conclude that interpretability provided by rule-based scales play an important role in understanding and accepting a CDSS output, especially when interpretation is done on the feature basis. The system's output is convincing, and the doctors can act upon it. Interpretations also help doctors to identify incorrect conclusions when a system produces them. They still see room for improvements, as not everything should be measured in numbers. The doctors see the importance of combining data-driven output with rule-based scales. Despite the understanding that data-driven models are based on high-quality, real-world data, doctors still ask for standard and known tools as they are created from the formal research results and widely accepted clinical guidelines. Doctors still believe that experts should be involved in the model development. This can potentially help to

expose the results of the CDSS even to patients. Our results show that a hybrid approach when a data-driven models are complimented with standard rule-based scales increases its acceptance and usefulness.

Acknowledgments

References

[1]   Wilson EV, Lankton N. Effects of Prior Use, Intention, and Habit on IT Continuance Across Sporadic Use and Frequent Use Conditions. Communications of the Association for Information Systems [Internet]. 2013 Sep 1 [cited 2021 Jul 20];33(1):3. Available from: https://aisel.aisnet.org/cais/vol33/iss1/3

[2]   Ahmad MA, Teredesai A, Eckert C. Interpretable machine learning in healthcare. Proceedings - 2018 IEEE International Conference on Healthcare Informatics, ICHI 2018. 2018 Jul 24;447.

[3]   de Clercq PA, Blom JA, Korsten HHM, Hasman A. Approaches for creating computer-interpretable guidelines that facilitate decision support. Artificial Intelligence in Medicine. 2004 May;31(1):1–27.

[4]   E C. Data-driven clinical decision processes: it's time. Journal of translational medicine [Internet]. 2019 Feb 12 [cited 2021 Aug 3];17(1). Available from: https://pubmed.ncbi.nlm.nih.gov/30755218/

[5]   Derevitskii I, Funkner A, Metsker O, Kovalchuk S. Graph-Based Predictive Modelling of Chronic Disease Development: Type 2 DM Case Study. Studies in health technology and informatics. 2019;261:150–5.

[6]   Kovalchuk SV, Kopanitsa GD, Derevitskii IV, Savitskaya DA. Three-stage intelligent support of clinical decision making for higher trust, validity, and explainability. 2020 Jul 25 [cited 2021 Aug 6]; Available from: https://arxiv.org/abs/2007.12870

[7]   Elkhovskaya L, Kabyshev M, Funkner A, Balakhontceva M, Fonin V, Kovalchuk S. Personalized Assistance for Patients with Chronic Diseases Through Multi-Level Distributed Healthcare Process Assessment. Studies in health technology and informatics. 2019;261:309–12.

[8]   Kahl KG, Schweiger U, Correll C, Müller C, Busch ML, Bauer M, Schwarz P. Depression, anxiety disorders, and metabolic syndrome in a population at risk for type 2 diabetes mellitus. Brain Behav. 2015 Mar;5(3):e00306

[9]   Lankton NK, Wilson EV, Mao E. Antecedents and determinants of information technology habit. Information and Management. 2010;47(5–6).

[10]   Ahlin E. Semi-Structured Interviews With Expert Practitioners: Their Validity and Significant Contribution to Translational Research. Semi-Structured Interviews With Expert Practitioners: Their Validity and Significant Contribution to Translational Research. 2019 Jan 3;