The Development and Feasibility of a Triage System for Use in Primary Care

1st Magnús Friðrik Helgason
Department of Computer Science
Reykjavik University
Reykjavik, Iceland
magnush19@ru.is

2nd Kári Steinn Hlífarson
Department of Computer Science
Reykjavik University
Reykjavik, Iceland
kari19@ru.is

3rd Baldur Olsen
Department of Computer Science
Reykjavik University
Reykjavik, Iceland
balduro19@ru.is

4th Steindór Oddur Ellertsson
Primary Health Care
of the Capital Area
Reykjavik, Iceland
steindor.ellertsson@gmail.com

5th Hrafn Loftsson
Department of Computer Science
Reykjavik University
Reykjavik, Iceland
hrafn@ru.is

6th Stefán Ólafsson
Department of Computer Science
Reykjavik University
Reykjavik, Iceland
stefanola@ru.is

***Abstract*—The Icelandic healthcare system faces challenges common to many Western countries, e.g., staff shortages and high workloads, with negative impacts on patient care. Long waiting times, overcrowded clinics, and overworked healthcare professionals contribute to diminished healthcare quality. To address these issues, we propose a system for patients and providers to alleviate pressure on primary care clinics. The system consists of an AI model for triage, a Questionnaire User Interface (QUI), and a Consultation User Interface (CUI). The QUI aims to reduce unnecessary in-person consultations by allowing patients to answer questions related to their well-being and medical history, similar to those posed during health center visits. The AI triage model assesses risk levels for specific disease groups by analyzing patient responses to the questions and determines the necessity of an in-person visit. The CUI facilitates the diagnosis and scoring of patients when further examination is required. The obtained risk scores inform treatment advice, optimizing healthcare professionals' decision-making. A feasibility study showed that the system could be integrated with the clinical workflow, such as automatically filling clinical text notes, saving doctors valuable time. By taking an user centered design approach and using AI for symptom scoring and patient classification, the system has the potential to enhance healthcare provider efficiency and improve consistency in patient care.**

***Index Terms*—triage system, user centered design, feasibility study, primary care**

I. Introduction

The health care system in Iceland, like in many other Western countries, faces significant challenges due to the lack of staff and high workload. Under-staffing in the healthcare system leads to lower-quality services and has a negative impact on patient outcomes. There is a high demand for medical services, far more than the health system staff can provide. This has led to long waiting times at health centers and overwhelming traffic in emergency rooms. Doctors are increasingly looking abroad for work in the hope of better working conditions, further exacerbating the problem [1], [2].

Patients often have to wait for hours in overcrowded clinics before they can see a doctor. Furthermore, when a specialist is needed, the wait for an appointment can be days, weeks, or even months. This has the effect that people are less likely to seek medical attention, despite the need for it, which can have serious, harmful consequences for their health. On the other hand, it is not uncommon for people to seek health care for problems that could be treated at home with self-treatment [3].

Increased workload and stress also affect healthcare professionals, who are forced to work faster and more efficiently to meet this increased demand [4]. This can have a serious effect on them, both physically and mentally, which can lead to misdiagnoses, mistakes, and burnout [5]. In addition, the increasing workload of doctors and nurses reduces their time with each patient. Physicians under high workload are also more likely to give in to pressure from patients and prescribe antibiotics or refer the patient for diagnostic tests that are not in accordance with clinical guidelines. Antibiotics have negative side effects and spur the creation of antibiotic-resistant bacteria that are estimated to cause the death of 5 million people annually worldwide [6].

The lack of staff in health centers and the high demand for medical assistance are not the only reasons why doctors have little time with patients. Outdated, inefficient, and unfriendly software – used by doctors to manage patient information, note visits, etc. – takes valuable time out of their day that could be better spent diagnosing and treating patients. Doctors in clinics have to spend three hours typing and processing data on a computer for every hour they spend with a patient [7]. Therefore, there is a need for solutions that can ease the work of healthcare professionals, reduce stress and increase their time with patients.

This work was supported by the Icelandic Student Innovation Fund.

In this paper, we present a system intended to reduce the pressure on primary care clinics. Electronic triage systems in healthcare have been most prevalent in the emergency medical services field. Their feasibility and applicability in a primary care setting are understudied and have mixed results [8], [9]. Our work describes the integration of artificial intelligence (AI), database, application programming interface (API), and user interfaces in a system for triaging patients seeking primary care. Furthermore, we present an evaluation of the feasibility of our approach.

The system is a web application and consists of a Questionnaire User Interface (QUI), a Consultation User Interface (CUI), and an API connecting system components, such as the AI triage model and the database. The main purpose of the system is to reduce the burden on healthcare workers due to high numbers of patients seeking primary care. The QUI is publicly accessible to individuals who intend to seek medical assistance at primary care clinics. These potential patients answer questions using this interface related to their well-being, current medical conditions, and their clinical history. The final set of questions that users get changes dynamically according to their symptoms and demographics. These questions are stored in a database accessible via an API.

The AI triage[1] model predicts risk levels (probability) of specific disease groups, based on patient responses to selected clinical questions, similar to those asked when patients seek medical care at the health centers. These risk scores are used to characterize patients by determining the likelihood of certain clinical disorders, such as strep throat or respiratory infections. By triaging patients in this manner before in-person consultations, low-risk patients can be offered other means of communication. This AI model was developed by Ellertsson et al. [10].

The CUI has modules for clinicians to maintain and structure the system's data, as well as to visualize and interpret research statistics related to their patients. Once a patient has answered the questionnaire, and it has been determined that their condition requires further examination by a healthcare professional, the system gathers all relevant information and uses the AI model to diagnose and score the patient. The risk scores obtained from this analysis are then displayed in the CUI and used to provide advice regarding the patient's treatment, such as recommending a blood test or prescribing antibiotics.

In Iceland, a clinical text note (CTN) accompanies all communications with a doctor, where relevant information about the patient's visit and diagnosis is recorded. Using the patient's responses, the consultation interface can automatically fill in these notes, saving doctors a lot of time and allowing them to focus on more important aspects of their practice. In addition, the use of AI in symptom scoring and classification of patients has the potential to relieve the burden on healthcare providers. By using AI to assign a risk score to patients, healthcare providers will be able to more easily assess which patients have more serious health problems and need quicker action. Finally, with sufficient data collection with the software in the coming years, this solution could lead to a certain baseline and consistency in the healthcare system in terms of treating patients with similar answers to the questionnaire. It could, for example, lead to better consistency in doctors' approaches when faced with similar medical problems.

[1]Triage is the preliminary assessment of patients in order to determine the urgency of their need for treatment and the nature of treatment required.

The rest of this paper is organized as follows: In Section II, we discuss related work, and in Section III, the design and the development of our system is presented. The purpose and execution of our feasibility study is presented in Section IV, along with its results in Section V. We discuss our findings in Section VI, and, finally, we conclude in Section VII.

II. Background

The development and use of electronic triage systems in healthcare has been most noticeable in emergency medical services, where triage aims to predict the severity of patient problems, with the aim of organising the patient flow [11]–[13].

However, the clinical efficacy of existing systems is not clear. Lidal et al. [11] concluded that there is a lack of evidence regarding the effectiveness of pre-hospital triage systems and their effects are not clear, particularly when a system is to be used in more than one setting. Additionally, Azerado et al. [12] conducted a systematic review of research on the use of the Manchester Triage System in the emergency room. This established system has proven validity for use in children, adults, patients with coronary syndrome and patients with acute pulmonary embolism. Results from 22 scientific papers indicated that the system was inclusive and had short term predictive power for emergency department admission and death.

Conversely, Dugas et al. [13] found that a computer-based electronic triage system (ETS) distributed patients more evenly across severity levels compared to the more commonly used Emergency Service Index (ESI) triaging approach. In a cross-sectional study of approximately 25000 adult emergency room patients, the ETS bases its patient distributions on the frequency of critical outcomes compared to the ESI, which focuses on resources utilized and critical outcomes.

The design of interfaces for healthcare workers has implications for data quality. A recent review of Electronic Health Records (EHR) interfaces showed that supporting users with clear and intuitive interface widgets and functionality, such as having mandatory fields, scaffolding interactions using templates, and contextually-aware auto-complete, improved data completeness and correctness [14]. A review of EHR systems showed that the usability problems tend to arise when interfaces violate natural dialog, lack consistency, have ineffective language use and information presentation, lack customization, clear feedback, and error prevention, and increase cognitive load [15]. Studies focusing on EHR system safety made no objective assessments and applied only inductive reasoning methods for hazard recognition. The review found that the implications of these usability problems for safety are inconclusive, though it is largely accepted in the literature that adhering to design guidelines and principles, e.g., Nielsen's usability heuristics[2], ultimately increases patient safety.

Some studies focus on the effectiveness and accuracy of the AI models themselves, as opposed to the use of the system or interfaces. An AI virtual assistant which provides patients with triage and diagnostic information was developed in [16]. Their findings showed that the AI system was able to provide patients with triage and diagnostic information with a similar accuracy to that of human doctors.

A triage pipeline that examines chest radiographs for the presence, severity, and progression of COVID-19 pneumonia was developed in [17]. The authors reported a diagnostic accuracy of 95% when evaluating the pipeline on a prospective cohort of 80 radiographs.

In primary care specifically, there are several examples of research in the literature on electronic and AI-driven triage systems. An investigation into the patterns of use of the online triage system `askmyGP` revealed that young people were the primary users, they found it easy to use, and that knowledge of the user's context is key to successful triage, whether by human or machine [8]. Another study on the use of digital communication in the Swedish healthcare system showed, through a qualitative analysis, that such systems can be cumbersome and unwieldy but worked well as a distributional and analytical tool for questionnaires, and that health care workers using these tools tended to make better use of their time [9]. Though the results were generally positive, it is difficult to generalize them to situations beyond those of these studies and more research in this area is needed.

In addition to having mixed or inconclusive results on the use of automatic methods for triage in primary care, it can be difficult to use human or computational methods to evaluate the ability of AI models (machine learning models, specifically) for automatic primary care triage. In an observational study using retrospective data, the inter-rater reliability (IRR) between AI models and general practitioners in primary care was explored [18]. The results showed that the IRR between general practitioners was too low for a reliable evaluation of the efficacy of AI models for primary care triage.

The work described in this paper focused on the design and development of a system for automatic triage of individuals seeking primary care. Fully incorporating digital applications into the health care workers daily life (i.e. domestication) is a non-trivial undertaking. A qualitative study by Andersen [19] revealed that systems for healthcare that take a small-scale approach and user-centered design methodology, addressing the needs and context of the users, had potential for domestication. Though these findings are hard to generalize, they show that healthcare workers often use many different systems to carry out their work and the interoperability between these systems is often lacking. Additionally, there is a substantial cognitive load to using several systems on a daily basis and although healthcare workers can and do study software instruction manuals, clear instructions, tutorials, and scaffolding users during use is important.

[2]https://www.nngroup.com/articles/ten-usability-heuristics/

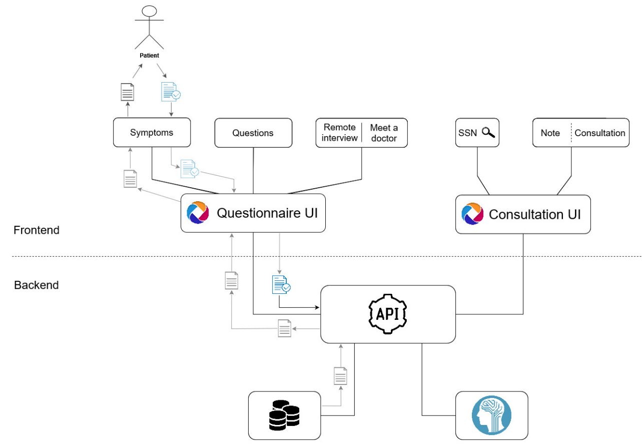

Fig. 1. Both the Questionnaire UI (QUI) and the Consultation UI connect to the same backend API, which interacts with the database and the AI model. The QUI presents patients with various symptoms, retrieved from the database through the API, that they have to choose from and prioritize. The patient's responses are then sent from the QUI to the API.

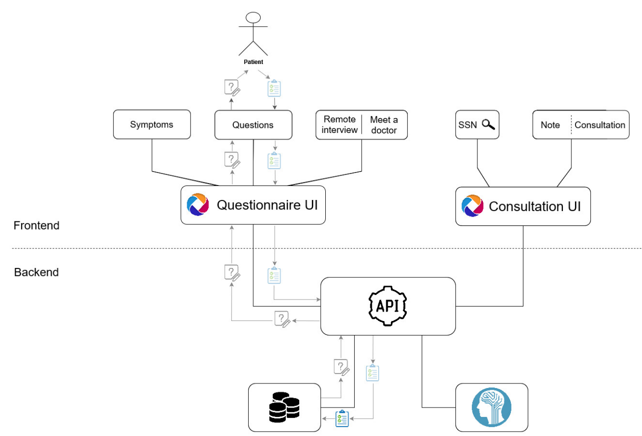

Fig. 2. The API queries the database to get a list of questions, specific to the symptoms given by the patient. Once the patient has finished answering the questionnaire, the answers are sent back to the API, which instructs the back end to save the answers in the database.

III. System Design and Development

The aim of our work is to create a system that makes it possible to use an AI model for triaging potential primary care patients into low or high risk groups. We developed two main interfaces, the Questionnaire User Interface (QUI) and the Consultation User Interface (CUI), as well as a secondary Administrator User Interface (AUI) to manage data and view usage statistics for the other parts of the system.

We developed a database that supports storage of question data and information necessary for managing data in the UIs, as well as an API allowing calls between the different components of the system.

The AI model used in our system was made for triaging patients with respiratory diseases before arrival to primary care clinics in Reykjavik, Iceland, and is described in previous work [10], [20]. The triage model consists of two main components: (1) a deep neural network for extracting clinical features from Clinician Text Notes (CTNs) [20]; and (2) a type of logistic regression that outputs a score between 0 and 1, where 1 means an increased probability of lower respiratory tract infection diagnosis [10]. The clinical feature extractor is based on a Transformer architecture (RoBERTa [21]) and pre-trained on the Icelandic Gigaword Corpus [22] to increase its capacity to process the Icelandic language. This model was trained to mark spans of text in CTNs that contain answers to a set of given clinical questions and was trained on notes with answer spans hand-annotated by clinicians and accompanying relevant questions. Given a set of questions, the feature extractor was then used to extract answer spans from a large set of CTNs that have particular diagnoses. The training objective for the triage model was predicting a patient's likelihood of having a lower respiratory tract infection. It was trained by using the clinical feature extractor on features that patients can be asked questions about through a web-based interface, such as the one presented in this paper, namely the questions presented by the QUI. Patients' responses to these questions are input to the triage model, which then outputs a likelihood score for lower respiratory tract infection.

The model was evaluated by using it to triage patients with respiratory diseases before coming to the primary care clinic. A follow-up was then conducted by looking at the patients' actual clinical outcomes after their visit to primary care and compared to how the model had triaged patients. The results showed that the model triaged all patients, that actually had a high risk of severe respiratory disease, correctly into the top five risk categories, while patients with a low risk ranked in the bottom five. Using this technology for primary care screening opens up the possibility of treating 35% of patients with respiratory symptoms at home, as well as reducing the number of antibiotic prescriptions by 25% and referrals for lung x-rays by 30% [10].

The front end of the system, developed using the React framework, is hosted on a Apache2 server in Google Cloud. The back end, which also runs in Google Cloud, uses the Django Rest Framework and an SQLite database.

A. UI Development

The QUI presents patients with various symptoms that they have to choose from and prioritize. These symptoms are retrieved from the database, through the API, and are the same for all users. The patient's responses, the selected symptoms and their priorities, are then sent from the QUI to the API (see Figure 1).

Using information from the user responses, the API queries the database to get a list of questions, specific to the given symptoms, to be answered by the patient in the QUI. Once the patient has finished answering the questionnaire, the answers are sent back to the API, which instructs the back end to save

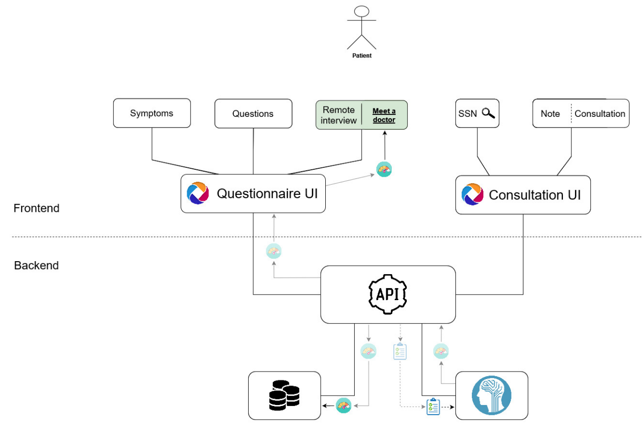

Fig. 3. The API sends the patient' answers to an AI model that calculates the risk score. The scores are saved in the database and sent to the QUI so that it can generate an appropriate response for the user based on the scores.

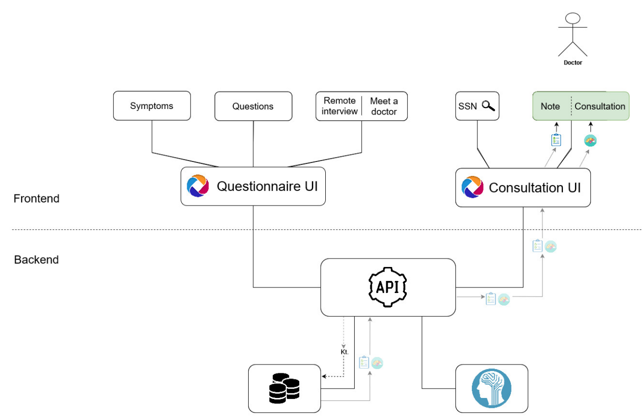

Fig. 4. A doctor finds the most recent answers to the questionnaire and the risk scores derived from them. This information is then used to create a Clinical Text Note (CTN) for the doctor to speed up the workflow.

the answers in the database. This process can be seen in Figure 2.

Once the answers are saved in the database, the API sends the answers to the AI model that calculates the risk score. The API then instructs the back end to save the scores in the database. Moreover, the API also sends the risk scores to the QUI so that it can generate an appropriate response for the user based on the scores (see Figure 3).

In order for the CUI system to be usable, the patient must have answered the questionnaire and be next in line in the queue for healthcare. Figure 4 show the flow in the system when the doctor selects one of the patients who chose, in the QUI, to see this doctor. The doctor finds the most recent answers to the questionnaire and the risk scores derived from them. This information is then used to create a CTN for the doctor to speed up the workflow. There are also future plans to use the risk levels to provide advice regarding patient treatment.

Fig. 5. Selection of symptoms

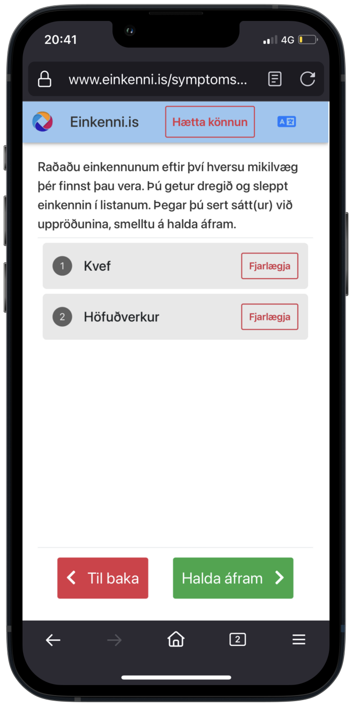

Fig. 6. Prioritization

Fig. 7. Opening question

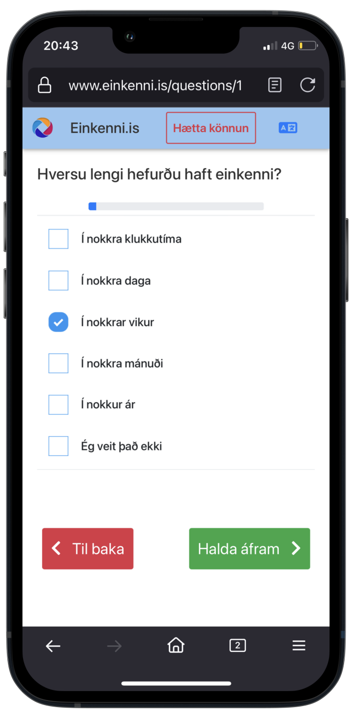

Fig. 8. Multiple choice question

Fig. 9. Input text question

Fig. 10. Yes/no question

The system as explained so far has been fully implemented for the process where a patient seeks help and a doctor provides a treatment. Given the modular architecture, further improvements of the system are possible. The latest addition to the system is the aforementioned AUI, which is divided into two parts: Statistics and Question Center. This interface simplifies system administrators to work with the database and read patient responses. A user of the CUI with administrator rights can access this part of the interface.

B. The Questionnaire User Interface

The QUI is intended for patients seeking medical care in health centers. Currently, the participants have been individuals in the waiting room at the clinic who take a few minutes while they wait to go through the questionnaire process for research purposes. The majority of these participants answered the questionnaire by using their smartphones. The QUI supports two languages, Icelandic and English. The first step before starting to answer the questionnaire is agreeing to the research terms. Next, the participants use their phone for electronic identification through https://island.is/. If authentication is successful, a research assistant gives users access to the web application via a secure authentication method. Finally, the participant is taken to the first step in the questioning process where they select symptoms.

After participants have gone through the research consent process, they are directed to another page where they select the main symptoms they have been experiencing. It is worth mentioning that from now on, users can press the "exit survey" button, which takes them back to the home screen and stops the process. Once a user has selected symptoms, the next step is to prioritize them by importance. After these steps, a user will be presented with tailored questions based on symptoms. The first question is always the same opening question, where the user is asked why they are seeking medical help. These steps can be seen in Figures 5–7.

From this point on, users answer customized clinical questions. The questions are divided into three categories: multiple choice questions, input text questions and yes/no questions (see Figures 8–10). Behind the scenes, questions are further divided into conditional, contingent, and dynamic questions. If a user answers certain questions positively, the list of questions can be updated to include new questions. An example of this can be seen in Figure 8 where users select a few weeks ("í nokkrar vikur"), but then they are asked in more detail (see Figure 9) how many weeks. The dynamic questions program takes in the most common symptoms that users did not check when ranking the symptoms, forms them as a multiple choice question, and then ask users again so that they do not miss important symptoms.

C. The Consulting User Interface

The CUI is intended for doctors to make full use of information obtained from patients who go through the QUI. Once doctors are logged in, they can put a social security number in the lookup search, or select patients that have an designated doctor, in order to see information regarding the patient's last visit via the QUI. The information includes automatically generated CTNs about the patient's visit, advice on treatment, and more.

Figure 11 shows the CTN, recommendations, red flags and other items that doctors can use to simplify the work process and rely on to better determine the patient's treatment. The recommendations at this point are not usable as the research process has not lasted long enough, but the design of the interface is ready to show what this tool will look like when the time comes. The CTN contains all the main information from the patient's answers. Doctors can then change and add to the note at will. They can also change or add certain values, such as temperature or blood pressure. When they press the "update advice" button, the new information will be added to the answers and the AI models will calculate new risk levels, resulting in a new CTN, recommendations, red flags, etc.

The latest addition to the CUI is the Administrator User Interface (AUI), which can be used by doctors with administrator rights. This interface can be opened by clicking on the user icon in the upper right corner of the page (see Figure 12). The AUI is composed of a statistics and a question center. The statistics interface is constantly evolving, but it provides an interpretation of all the most important aspects of patient use of the QUI. The interface shows answers, risk levels, dates and information for users who have used the system and those who are answering the questionnaire in real time. Doctors who have access to this system can record the results of patients who have completed the questionnaire and view statistics on those who have had their results recorded. These statistics show the amount and percentage of outcomes according to risk scores, which are divided into two tables: the risk level – divided into two intervals with a selected transition point separating them – and the risk score – divided into 10 intervals. Doctors can choose whether the outcomes are active or inactive and the statistics will only be based on active results. The AUI can be seen in Figure 12.

The Question Hub (QHub) is the other side of the AUI. It gives the administrator a better overview of the questions that are in the system database and makes it easy to edit them. The admin can view, delete, change and add questions, put questions into groups that suit the research project, such as making groups of questions that relate to each AI model. Questions can be searched for or filtered by question type, which makes it easier for doctors to get the information they are looking for.

IV. Feasibility Study

We assessed the feasibility of our main user interfaces: the QUI and CUI. The feasibility of the QUI was assessed among patients waiting for an appointment with a doctor in a primary care setting. First, participants were asked to consent to participating in our study. Next, participants answered questions in the QUI that led them through the process described in Section III-B. Patient responses were recorded, as was the time it took them to use the interface. Three questions were inserted into the interface specifically for the feasibility study to get feedback regarding the user experience. These questions were Likert scale items, with anchors corresponding to levels of

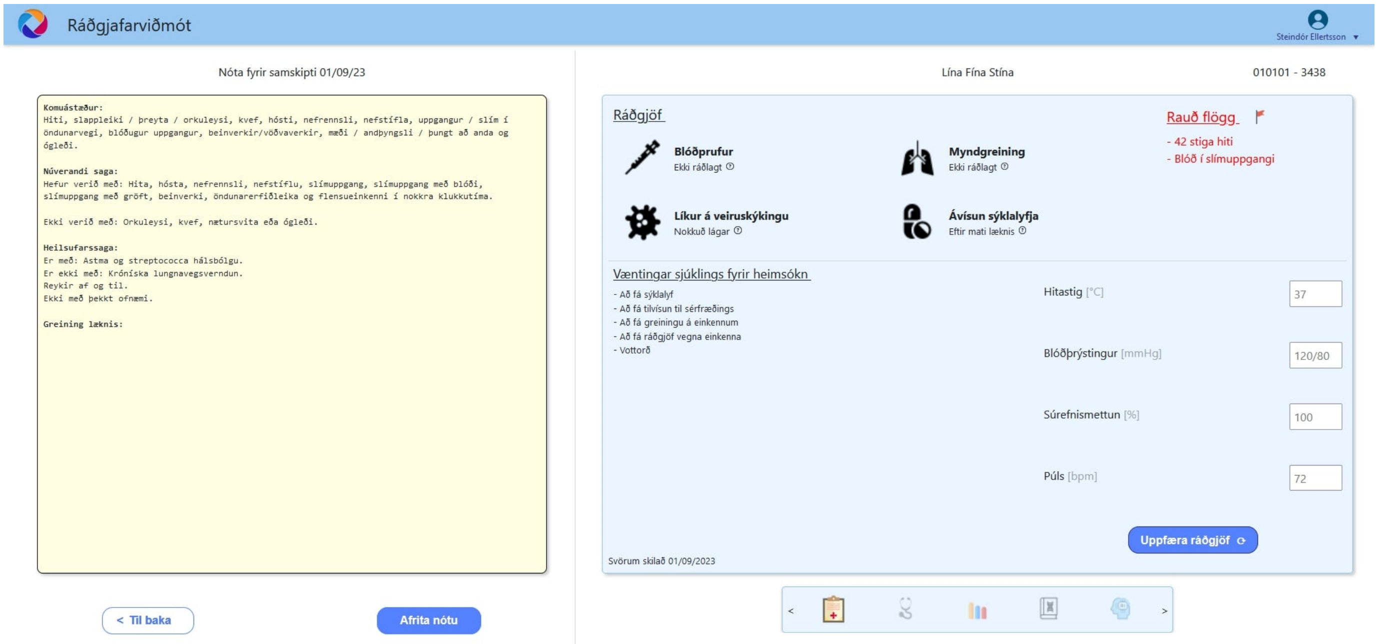

Fig. 11. A clinical text note (CTN) and the consultation user interface (CUI). The CTN contains all the main information from the patient's answers. The doctor can then change and add to the note at will. A doctor can also change or add certain values, such as temperature or blood pressure. When a doctor presses the "update advice" button, the new information will be added to the answers and the AI model will calculate new risk levels, resulting in a new CTN, recommendations, red flags, etc.

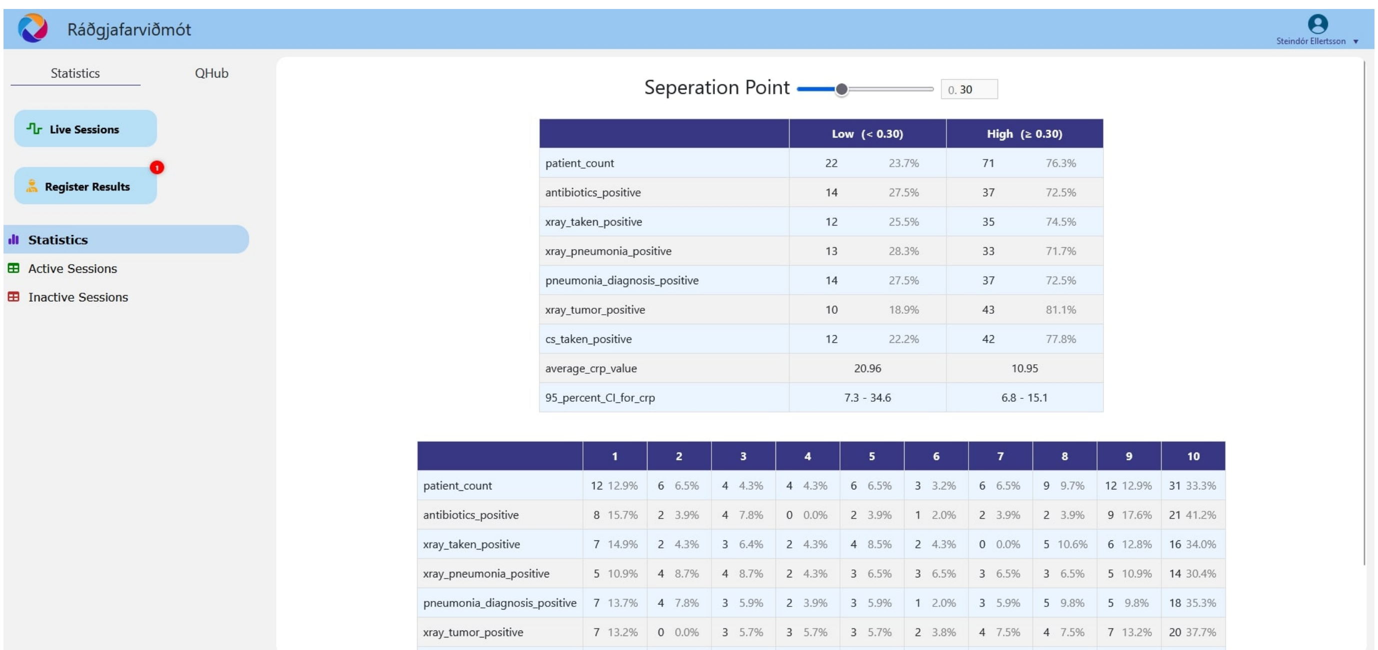

Fig. 12. The administration user interface (AUI) shows answers, risk levels, dates and information for users who have used the system and those who are answering the questionnaire in real time.

agreement. These questions can be seen in the results in Table I and Figure 13.

For the feasibility study of the CUI, we tried to get as much feedback as possible from each participant, as it is difficult to get busy healthcare professionals to participate. In the CUI, the doctor receives an automatically generated CTN and recommendations regarding patient treatment. In order to assess the feasibility of this interface, a think-aloud protocol was conducted. It involved giving participants, who have no knowledge of the system's internals, specific tasks to solve

TABLE I
THE RESULTS FROM PARTICIPANT RATINGS OF THE QUESTIONNAIRE USER INTERFACE, THE MEDIAN AND INTER-QUARTILE RANGE. EACH QUESTION WAS A LIKERT-SCALE ITEM RANGING FROM 'STRONGLY DISAGREE' TO 'STRONGLY AGREE', WITH VALUES FROM 1 TO 5.

Question	**Median (IQR)**
I found this website easy to use	5 (1)
I would like to use this website again	4 (1)
This website looks professional	4 (1)

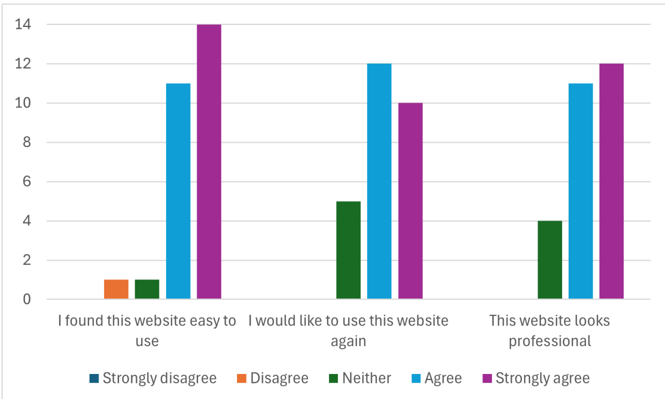

Fig. 13. The frequency of the type of responses to the Likert-scale items concerning the questionnaire interface.

and they were asked to express their thoughts as they carried out the tasks. The time it took the participants to solve these tasks was also recorded.

V. RESULTS

In this section, we present the results of the feasibility studies for both the QUI and the CUI.

A total of 27 patients participated and provided insight into their health, the time it took to complete the questionnaire, and their overall satisfaction with the user experience. Of these 27 participants, 8 were male and 19 were female. The average age of the participants was 49 (18.03 standard deviation), with the youngest being 19 and the oldest 81.

The average time it took participants to complete the questionnaire was 7 minutes and 58 seconds, with a standard deviation of 3 minutes and 39 seconds.

Table I summarizes the ratings participants gave to the QUI on three Likert-scale items. For the question "I found this website easy to use" (see Figure 13), 25 out of the 27 participants said they either agreed or strongly agreed that the website was easy to use. Responses to "I would like to use this website again" show that 22 participants are inclined towards the positive answer options and stated that they would like to use this website again. For the question "This website looked professional", 23 participants either agreed or strongly agreed that the website looked professional.

Two individuals participated in the interviews and the think-aloud testing for the CUI. Both individuals have experience in health care, as one being a seasoned primary care doctor and the other a senior medical student. The benchmark time set

Task	Time taken by participant 1 [minutes]	Time taken by participant 2 [minutes]
Log in (predefined credentials)	Completed	Completed
Find patient with given SSN	Completed	Completed
Write something in "Skoðun" in the CTN	Completed	Completed
Review the red flags	Completed	Completed
Review the consultation	Completed	Completed
Explain the reason behind the patients visit	Completed	Completed
Find the patients name	Completed	Completed
What date did he take the questionnaire?	Completed	Completed
Copy the Note to your clipboard	Completed	Completed
What does the patient expect from the visit?	Completed	Completed
Total time [minutes]	4:21	5:34

Fig. 14. Results from the tasks that the individuals had to solve.

by the team was three minutes. The tasks that the individuals had to solve can be seen in Figure 14.

Participants were asked to think out loud as they interacted with the prototype and were asked several questions about their experience using the CUI, if there was anything missing, and if they found it useful in a semi-structured interview. Using thematic analysis, we identified three major themes. The participants expressed *enjoyment* using the interface and anticipation for how it could be used. The second theme was *usefulness*, as they found the automatically generated CTNs useful and saw how it could shorten and simplify doctors' work processes significantly. The third major theme identified by both interviewees was *trust*. They expressed that it would be difficult to convince healthcare workers and providers that AI-based advisors were reliable, and that thorough and convincing research needs to be carried out to demonstrate their usefulness.

VI. DISCUSSION

The aim of this project was to develop a system that uses two user interfaces and an AI model to facilitate, shorten and refine the time-consuming process that takes place when patients seek medical assistance in Iceland's primary care facilities. This is a difficult challenge, but necessary if the quality of life of individuals and health professionals in the country is to be improved. According to our results, this system will give patients an easier and faster way to receive symptom diagnosis, and even clinical indications for treatment, and reduce doctors' time in computer and paper processing, freeing time for them to better care for patients.

The standard deviation of the mean time it took patients to complete the QUI tells us that there was a very wide range of time around the mean. Such variability could be attributed to different factors such as the users' knowledge of digital interfaces, the clarity of the questions or the users' state of health when they answer. From Table I and Figure 13, it can be concluded that users generally found the website easy to use. However, there were very few who either disagreed or had a neutral assessment. This suggests that although the majority had a positive experience, there is room for improvement for making the interface even more user-friendly. The responses of the participants regarding whether they wanted to use the website again showed mixed results (see Figure 13). A few users wanted to use the website again, but half checked "neither agree nor disagree" as an answer. To some extent, this neutral reaction can be attributed to people being in situations they would not normally want to be in (i.e. sick in a queue at health care) and so it is likely that some people would include it in their thinking with regard to whether they would like to answer the list of questions again. Regarding the professional appearance of the website (see Figure 13), the majority of users seemed to have a positive attitude, with a noticeable tendency to "agree" and "strongly agree". This is a strong indication that the design, appearance and content of the website are in line with user expectations.

In the feasibility study for the CUI, both individuals managed to complete their tasks quickly. Even though both individuals exceeded the set time limit of three minutes, the team is happy with the results, with most of the time spent on comments and brainstorming. The individuals' first reaction to seeing certain parts of the system, such as a CTN or advice, was often to analyze them in detail before continuing. This enthusiasm on the part of the participants resulted in a lot of good comments and constructive criticism, but also resulted in more time to go through the study. Neither of the individuals needed assistance in solving the tasks, which indicates that the system is easy to use.

It is worth noting the limitations of the feasibility studies. The main limitation is the number of participants in the studies, especially the CUI where only two people participated. 27 people took part in the survey regarding the QUI. This number is sufficient to give a rough idea of the time user spend answering questions about their symptoms and people's attitude towards the system. Nevertheless, a larger population is needed for more accurate and reliable results. Moreover, a much larger population is needed to have any chance of meaningful results, such as correlation tests.

A further limitation is the circumstances of the feasibility study of the QUI, where being ill and taking the survey in a healthcare queue could influence user responses. For example, it is likely that a person with a mild cold who takes a survey as soon as they arrive at the health center will have a generally more positive attitude than a person with a temperature of over 40 degrees who has been waiting for an hour for medical assistance.

VII. Conclusion

The challenges faced by the healthcare system in Iceland, such as understaffing, high workload, and inefficient software, are contributing to a decline in the quality of services and negative impacts on patient outcomes. The increasing demand for medical services has resulted in increased waiting times for primary care services. Healthcare professionals, facing heightened workload and stress, are more prone to errors and misdiagnoses, with potential consequences for patient safety.

The presented system, consisting of a Questionnaire User Interface (QUI), Consultation User Interface (CUI), and an AI triage model, offers a promising solution to alleviate the burden on primary care clinics. By allowing patients to provide relevant information through the QUI and employing the AI triage model to predict risk levels of specific disease groups, the system enables more efficient and targeted patient management. According to results from our feasibility studies, the QUI gives patients an easy and accessible way to receive symptom diagnosis, and even clinical indications for treatment. Moreover, it can reduce doctors' time in computer and paper processing, freeing time for them to better care for patients.

The CUI streamlines the workflow for healthcare professionals, automating clinical text notes (CTN) and providing valuable insights through research statistics. This not only saves time but also enhances the accuracy and consistency of patient records.

Future evaluations should assess the system's accessibility, navigability, and clarity of concepts, particularly for populations with specific accessibility needs, such as senior citizens and individuals with physical impairments. Additionally, future trials should evaluate the clinical reliability and effectiveness of electronic triage using the system across a larger and more diverse patient population, as well as types of diseases.

The incorporation of AI in symptom scoring and patient classification has the potential to enhance the overall effectiveness of healthcare providers, enabling quicker identification of patients with urgent healthcare needs. As the system collects more data over time, it has the potential to establish a baseline and promote consistency in medical approaches for patients with similar conditions.

References

[1] G. A. Gunnarsdóttir, "Yfir 600 íslenskir læknar í útlöndum [Over 600 Icelandic doctors abroad]," *Læknablaðið [The Icelandic Medical Journal]*, vol. 107, no. 1, 2021. [Online]. Available: https://www.laeknabladid.is/tolublod/2021/01/nr/7587

[2] M. Ladan, "The doctor exodus – with focus on the Scandinavian context," *BMJ Global Health*, 2023, blog. [Online]. Available: https://blogs.bmj.com/bmjgh/2023/04/13/the-doctor-exodus-with-focus-on-the-scandinavian-context/

[3] S. Renati and J. A. Linder, "Necessity of office visits for acute respiratory infections in primary care," *Family Practice*, vol. 33, no. 3, pp. 312–317, Jun. 2016.

[4] F. D. R. Hobbs, C. Bankhead, T. Mukhtar, S. Stevens, R. Perera-Salazar, T. Holt, and C. Salisbury, "Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14," *The Lancet*, vol. 387, no. 10035, pp. 2323–2330, Jun. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140673616006206

[5] Medscape, "'I Cry but No One Cares': Physician Burnout & Depression Report 2023," 2023. [Online]. Available: https://www.medscape.com/slideshow/2023-lifestyle-burnout-6016058

[6] C. J. L. Murray, K. S. Ikuta, F. Sharara, L. Swetschinski, G. R. Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, S. C. Johnson, A. J. Browne, M. G. Chipeta, F. Fell, S. Hackett, G. Haines-Woodhouse, B. H. K. Hamadani, E. A. P. Kumaran, B. McManigal, S. Achalapong, R. Agarwal, S. Akech, S. Albertson, J. Amuasi, J. Andrews, A. Aravkin, E. Ashley, F.-X. Babin, F. Bailey, S. Baker, B. Basnyat, A. Bekker,

R. Bender, J. A. Berkley, A. Bethou, J. Bielicki, S. Boonkasidecha, J. Bukosia, C. Carvalheiro, C. Castañeda-Orjuela, V. Chansamouth, S. Chaurasia, S. Chiurchiù, F. Chowdhury, R. C. Donatien, A. J. Cook, B. Cooper, T. R. Cressey, E. Criollo-Mora, M. Cunningham, S. Darboe, N. P. J. Day, M. D. Luca, K. Dokova, A. Dramowski, S. J. Dunachie, T. D. Bich, T. Eckmanns, D. Eibach, A. Emami, N. Feasey, N. Fisher-Pearson, K. Forrest, C. Garcia, D. Garrett, P. Gastmeier, A. Z. Giref, R. C. Greer, V. Gupta, S. Haller, A. Haselbeck, S. I. Hay, M. Holm, S. Hopkins, Y. Hsia, K. C. Iregbu, J. Jacobs, D. Jarovsky, F. Javanmardi, A. W. J. Jenney, M. Khorana, S. Khusuwan, N. Kissoon, E. Kobeissi, T. Kostyanev, F. Krapp, R. Krumkamp, A. Kumar, H. H. Kyu, C. Lim, K. Lim, D. Limmathurotsakul, M. J. Loftus, M. Lunn, J. Ma, A. Manoharan, F. Marks, J. May, M. Mayxay, N. Mturi, T. Munera-Huertas, P. Musicha, L. A. Musila, M. M. Mussi-Pinhata, R. N. Naidu, T. Nakamura, R. Nanavati, S. Nangia, P. Newton, C. Ngoun, A. Novotney, D. Nwakanma, C. W. Obiero, T. J. Ochoa, A. Olivas-Martinez, P. Olliaro, E. Ooko, E. Ortiz-Brizuela, P. Ounchanum, G. D. Pak, J. L. Paredes, A. Y. Peleg, C. Perrone, T. Phe, K. Phommasone, N. Plakkal, A. Ponce-de Leon, M. Raad, T. Ramdin, S. Rattanavong, A. Riddell, T. Roberts, J. V. Robotham, A. Roca, V. D. Rosenthal, K. E. Rudd, N. Russell, H. S. Sader, W. Saengchan, J. Schnall, J. A. G. Scott, S. Seekaew, M. Sharland, M. Shivamallappa, J. Sifuentes-Osornio, A. J. Simpson, N. Steenkeste, A. J. Stewardson, T. Stoeva, N. Tasak, A. Thaiprakong, G. Thwaites, C. Tigoi, C. Turner, P. Turner, H. R. v. Doorn, S. Velaphi, A. Vongpradith, M. Vongsouvath, H. Vu, T. Walsh, J. L. Walson, S. Waner, T. Wangrangsimakul, P. Wannapinij, T. Wozniak, T. E. M. W. Y. Sharma, K. C. Yu, P. Zheng, B. Sartorius, A. D. Lopez, A. Stergachis, C. Moore, C. Dolecek, and M. Naghavi, "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis," *The Lancet*, vol. 399, no. 10325, pp. 629–655, Feb. 2022, publisher: Elsevier. [Online]. Available: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)02724-0/fulltext?ref=healthtips.kr

[7] C. Sinsky, L. Colligan, L. Li, M. Prgomet, S. Reynolds, L. Goeders, J. Westbrook, M. Tutty, and G. Blike, "Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties," *Annals of Internal Medicine*, vol. 165, no. 11, pp. 753–760, 2016.

[8] A. Eccles, M. Hopper, A. Turk, and H. Atherton, "Patient use of an online triage platform: a mixed-methods retrospective exploration in uk primary care," *British Journal of General Practice*, vol. 69, no. 682, pp. e336–e344, 2019.

[9] A. C. Eldh, A. Sverker, P. Bendtsen, and E. Nilsson, "Health care professionals' experience of a digital tool for patient exchange, anamnesis, and triage in primary care: qualitative study," *JMIR human factors*, vol. 7, no. 4, p. e21698, 2020.

[10] S. Ellertsson, H. D. Hlynsson, H. Loftsson, and E. L. Sigurðsson, "Triaging patients with artificial intelligence for respiratory symptoms in primary care to improve patient outcomes: A retrospective diagnostic accuracy study," *Annals of Family Medicine*, vol. 21, no. 3, pp. 240–248, 2021.

[11] I. B. Lidal, H. H. Holte, and G. E. Vist, "Triage systems for pre-hospital emergency medical services – a systematic review," *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, vol. 21, no. 1, p. 28, 2013. [Online]. Available: https://doi.org/10.1186/1757-7241-21-28

[12] T. R. M. Azeredo, H. M. Guedes, R. A. Rebelo de Almeida, T. C. M. Chianca, and J. C. A. Martins, "Efficacy of the Manchester Triage System: a systematic review," *International Emergency Nursing*, vol. 23, no. 2, pp. 47–52, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1755599X14000512

[13] A. F. Dugas, T. D. Kirsch, M. Toerper, F. Korley, G. Yenokyan, D. France, D. Hager, and S. Levin, "An Electronic Emergency Triage System to Improve Patient Distribution by Critical Outcomes," *Journal of Emergency Medicine*, vol. 50, no. 6, pp. 910–918, 2016. [Online]. Available: https://doi.org/10.1186/1757-7241-21-28

[14] O. O. Madandola, R. I. Bjarnadottir, Y. Yao, M. Ansell, F. Dos Santos, H. Cho, K. Dunn Lopez, T. G. Macieira, and G. M. Keenan, "The relationship between electronic health records user interface features and data quality of patient clinical information: an integrative review," *Journal of the American Medical Informatics Association*, vol. 31, no. 1, pp. 240–255, 2024.

[15] M. Zahabi, D. B. Kaber, and M. Swangnetr, "Usability and safety in electronic medical records interface design: a review of recent literature and guideline formulation," *Human factors*, vol. 57, no. 5, pp. 805–834, 2015.

[16] A. Baker, Y. Perov, K. Middleton, J. Baxter, D. Mullarkey, D. Sangar, M. Butt, A. DoRosario, and S. Johri, "A Comparison of Artificial Intelligence and Human Doctors for the Purpose of Triage and Diagnosis," *Frontiers in Artificial Intelligence*, vol. 3, 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frai.2020.543405

[17] C. K. Kim, J. W. Choi, Z. Jiao, D. Wang, J. Wu, T. Y. Yi, K. C. Halsey, F. Eweje, T. M. L. Tran, C. Liu, R. Wang, J. Sollee, C. Hsieh, K. Chang, F.-X. Yang, R. Singh, J.-L. Ou, R. Y. Huang, C. Feng, M. D. Feldman, T. Liu, J. S. Gong, S. Lu, C. Eickhoff, X. Feng, I. Kamel, R. Sebro, M. K. Atalay, T. Healey, Y. Fan, W.-H. Liao, J. Wang, and H. X. Bai, "An automated COVID-19 triage pipeline using artificial intelligence based on chest radiographs and clinical data," *npj Digital Medicine*, vol. 5, no. 1, p. 5, Jan. 2022. [Online]. Available: https://doi.org/10.1038/s41746-021-00546-w

[18] A. Entezarjou, A.-K. E. Bonamy, S. Benjaminsson, P. Herman, and P. Midlöv, "Human- Versus Machine Learning–Based Triage Using Digitalized Patient Histories in Primary Care: Comparative Study," *JMIR medical informatics*, vol. 8, no. 9, p. e18930, 2020.

[19] S. Andersen, "eHealth Systems' User Experiences from Healthcare Workers," *Research in Social Sciences and Technology*, vol. 4, no. 1, pp. 51–72, 2019.

[20] H. Hlynsson, S. Ellertsson, J. Daðason, E. Sigurdsson, and H. Loftsson, "Semi-supervised Automated Clinical Coding Using International Classification of Diseases," in *Proceedings of the 5th International Conference on Natural Language and Speech Processing (ICNLSP 2022)*, M. Abbas and A. A. Freihat, Eds. Trento, Italy: Association for Computational Linguistics, 2022, pp. 95–106. [Online]. Available: https://aclanthology.org/2022.icnlsp-1.11

[21] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," *arXiv preprint arXiv:1907.11692*, 2019.

[22] S. Steingrímsson, S. Helgadóttir, E. Rögnvaldsson, S. Barkarson, and J. Guðnason, "Risamálheild: A Very Large Icelandic Text Corpus," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis, and T. Tokunaga, Eds. Miyazaki, Japan: European Language Resources Association (ELRA), May 2018. [Online]. Available: https://aclanthology.org/L18-1690