ICIC Express Letters
Part B: Applications
Volume **14**, Number **1**, January **2023**
 ISSN 2185-2766
pp. 1–11

SUPPORTING RURAL OUTPATIENTS THROUGH A WEB-BASED CONVERSATIONAL USER INTERFACE

Chakkrit Snae Namahoot[1,*] and Michael Brückner[2]

[1]Faculty of Science
[2]Naresuan University International College
Naresuan University
Phitsanulok 65000, Thailand
michaelb@nu.ac.th
*Corresponding author: chakkrits@nu.ac.th

Received March 2022; accepted June 2022

Abstract. *Even before the current (2020/2021) pandemic began, Conversational User Interfaces (CUI) had been seen as a valuable way to ease the burden on medical staff in many countries. In times of restricted direct contact with people, the need for online or virtual tools to connect patients with physicians has become even more visible. In addition, these restrictions hamper the training conditions for prospective doctors. This paper describes the design and implementation of a CUI covering patients' minor complaints of the Ear, Nose, and Throat (ENT), which can be correlated with infection by COVID-19. The purpose of this study is to provide pilot test results for an online anamnesis and diagnosis tool supporting the cooperative work of specialists and non-specialists at their workplaces. We have designed and created the cooperative online anamnesis and diagnosis system (COLDS) using 1) a knowledge-based system for the anamnesis mainly of complaints related to ENT including the eyes, 2) a knowledge base of disorders regarding ENT and eyes, and 3) a user interface that assists patients as well as cooperative processes involving non-specialists and specialists. COLDS is part of a clinical decision support system. The system has been evaluated in a two-tier pilot test process set in a real-life environment: Tier 1 was concerned with the usability of the system; whereas Tier 2 involved medical specialists to evaluate the outcome and recommendations created by the system based on an adapted Objective Structured Clinical Examination (OSCE) framework. Medical interns and doctors evaluated the system with a five-point Likert scale and the results show that 4.38 for the ease of system and 4.51 for overall satisfaction with the system at the confidence interval 95%.*
Keywords: Web-based diagnosis, Conversational user interface, Computer supported collaborative work, Artificial intelligence, Knowledge-based clustering

1. **Introduction.** Digital health care information related to individuals and populations is beginning to evolve in Thailand, since many institutions have employed Electronic Health Records (EHR). However, the data are currently only locally available. The patients' lack of access to their health data leads to less self-initiative in health care and less commitment to the care facility. Thailand's Community Health Centers are primarily small clinics that treat patients of all ages. Patients from simple ailments to life-threatening disorders make it difficult to anticipate how long the average patient will spend on a given day, and health workers frequently fall behind schedule. In remote areas, it becomes more and more common to initially use 'self-diagnosis' with the help of tools and smartphone apps that enable monitoring heart rate and blood pressure, among others. These tools generate valuable health data, which should be shared with the caretaker organization. Linking these data to the EHR accelerates the anamnesis process and thus

DOI: 10.24507/icicelb.14.01.1

leads to better utilization of the medical staff. This is the goal we aimed to achieve with COLDS.

Regarding the pandemic, current public health apps in Thailand mainly focus on tracking (Thai Chana) and on collecting symptoms after vaccinations (Mor Phrom) without any interactive components collecting and assessing symptoms. In the case of COLDS, the patients use a Conversational User Interface (CUI). Medical history and decision-making are the two main components of this process. The following sections will go through these elements in greater depth, focusing on conditions of the Ears, Nose and Throat (ENT), eyes, and the respiratory system. We focus on ENT because around 20% of adult patients in general practice consultations present otolaryngology conditions. Of those patients, around 30% report tonsillitis or severe sore throat and around 15% report persistent nasal symptoms; therefore, ENT is a reasonable target for telediagnosis as far as minor complaints are concerned. Ausayakhun et al. [1] have found that telemedicine can have a significant impact on health in rural areas, where ophthalmologists are rare. Web-based interaction is a promising development to engage patients with their medical institution.

Current CUIs in healthcare reportedly lack integration with general practitioners and other medical professionals [2,3]. Effective digital tools and smartphone apps are expected to increase satisfaction with healthcare services and better adherence to treatment guidelines provided by physicians directly to the patients. Moreover, better integration may well lead to higher customer loyalty, which would increase the organization's revenue as well. It is also important to promote trust in the patient-doctor relationship, which can certainly be achieved through the successful integration of the services. Therefore, we sought to design and develop a practical and safe Web-based diagnosis instrument, which connects patients at home with professional but still local, and thus familiar, contact persons.

A number of studies have addressed AI-assisted self-diagnosis. In the context of regional communities, we also have to provide links between an online system and experts to obtain the best of both worlds [4]. For the sake of brevity, we refer to [2] for helpful suggestions related to the CUI design: 1) Limit the intricacy of the interactions to those that a digital assistant might safely do, 2) schedule visits based on the severity of symptoms, 3) monitor health status, 4) remind and warn nurses when parameters become out of control, and 5) provide comfort and support until the next doctor's appointment. Furthermore, we must not overlook the importance of integrating the systems into the care process, which necessitates integration with eHealth and IT healthcare systems.

For the purpose of this paper, the user interface is in part translated to English language. The rest of the paper is structured as follows. Following a brief overview of previous work, we outline the design and system architecture of COLDS and its knowledge base, thereby laying emphasis on the development of the user interface. The next section reports on the extensive system testing and its results, which are discussed in more detail. Finally, we draw conclusions and point to important future aspects of this ongoing research.

2. **Requirements and System Design.** We have applied a user-centered design approach to building the user interface right from the onset of the project, which minimizes potential design issues related to the user interface [5]. In the course of writing the requirements for usability and accessibility of the system, we performed a user and task analysis to address the variations in roles and provide an appropriate balance of accessibility requirements [6]. Details are presented in the COLDS system architecture (Figure 1). In the following we give an overview of the general design considerations, which guided this research, present two of the numerous use scenarios, and detail the system architecture.

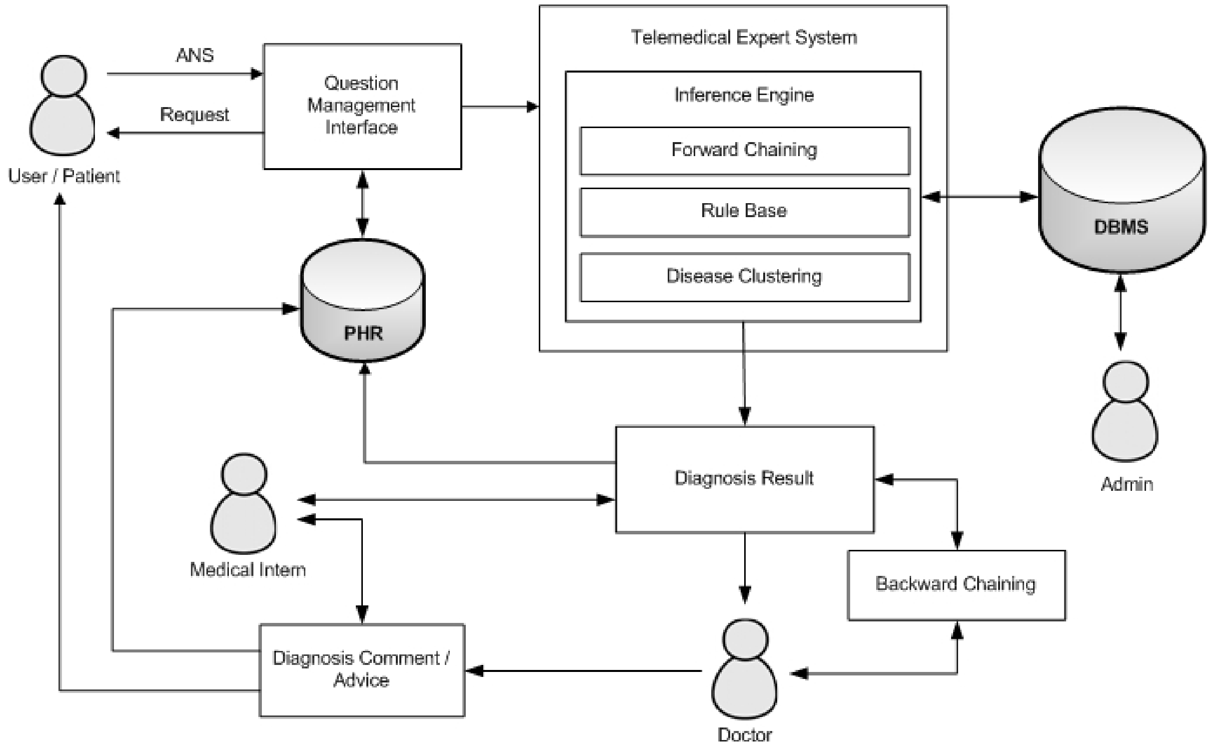

FIGURE 1. COLDS system architecture

To be effective, systems like COLDS need to support physicians with content that is relevant to the situation at the time of decision-making. This is mostly a matter of provision and effective presentation of appropriate data, which allows for quick and unambiguous interpretation. Since the workflow within COLDS is twofold (physician-oriented and automated), this basic requirement holds for both human interpretation and automatic processing of the EHR data and patients' responses to questions. The overall workflow is

```
Start the session
The patient has a chief complaint
The patient accesses COLDS UI via Web browser
     If the automated diagnosis is successful
             Physician checks diagnosis process
                     If automated diagnosis accepted
                             Send diagnosis to the patient
                     If automated diagnosis not accepted
                             Diagnosis process taken over by physician
                                     If diagnosis from data possible
                                             Send diagnosis to the patient
                                     If diagnosis from data not possible
                                             Let the patient see the doctor
     If automated diagnosis not successful
Diagnosis process taken over by physician
                     If diagnosis from data possible
                             Send diagnosis to the patient
                     If diagnosis from data not possible
                             Let the patient see the doctor
Terminate the session
```

In a recent review regarding interface design for user-centered decision support systems over the last three decades, Horsky et al. [7] have compiled a number of principles as best practices, which have been applied to the design of the user interface of COLDS. These best practices have guided the design process of COLDS and are summarized as follows:

- Consistency of design concepts, visual formats and terminology (wording) adapted to users (patients and medical staff): use simple and everyday medical terms, so that patients understand,
- Speedy interaction of the system with physicians at convenient points in the workflow (managing information density on the screen): automatic rule-based responses if possible and appropriate,
- Automatically processed results can be traced back by physicians (avoiding 'black box' advice): log the input and results derived by COLDS,
- COLDS system admins receive notifications of overridden diagnoses gained automatically and of diagnose process variables that did not lead to a successful diagnose process; these notifications may result in critical reviews of the system logic.

The COLDS system architecture consists of the following components (see Figure 1): Patient data are stored in the EHRs; most importantly, diagnosis results and comments are used for upcoming follow-up checks and treatments. The knowledge base comprises data prepared by collecting information about disorders, which contains such information as the name of signs, conditions, symptoms, disorders, treatments, complications, and recommendations. The Question Management Interface is used for the management of individual information on signs of disorders (COLDS focuses on ENT and eye healthcare) and the provision of appropriate closed questions to users/patients. Inference engine mechanisms: forward chaining is used for discovering the answer of a disorder that the user is suffering from while using COLDS. The mechanism examines all signs, symptoms and conditions from the answers provided by patients, e.g., after patients have stated their chief complaints, and the system requests responses to questions about their conditions. Patients must offer correct yes or no answers so that the system can return with additional questions based on past patient responses and attempt to identify the correct ailment with the fewest possible inquiries. Based on the knowledge base, the rule base was created as a disorder diagnosis decision tree for the ENT system and the eyes. The term "disease clustering" refers to the grouping of illnesses with similar or identical indications or symptoms. 1) Signs and symptoms are sorted and then associated by classification in the system architecture; clustering records any types of diseases and related symptoms; 2) this facilitates the user/patient answering useful inquiries, and 3) supply fast access to required information. The diagnosis result is shown after analyzing the data from user/patient specified/answer to the questions, e.g., the system checks all diagnosis questions and answers the patients have provided, and then the system finds and displays appropriate diagnosis results. COLDS employs backward chaining to allow clinicians to double-check and track the answers to questions; for example, if the right answer is chronic bronchitis, the system double-checks all questions and answers provided by the patients. In this method, doctors can assess the accuracy of the diagnosis and recommend treatment or additional health screenings. To present questions that patients have answered during the session, an answer revision procedure has been implemented. To avoid misdiagnosis, doctors give feedback through advice and approve or revise the diagnosis decision and comment and advise treatment of the patient. Also, medical interns can use this recommendation for case studies and practice.

The system has been set up to handle patient queries and present all possible diagnosis results, as well as the doctor's advice on the patients' situations. The steps in the procedure are as follows. First, patients have to access the Website and log in. Second, the system then enquires for signs and symptoms, uses the answers to ask further questions using question management and the system generates a diagnosis result with the assistance of the expert system. Questions asked by COLDS are supported by context-aware multimedia that help display body parts asked for in the anamnesis process; as an example, if asked for 'fever' a short on-demand video demonstrates how to measure body temperature with different tools.

The system sends the patient's information to the proper health professional/doctor after receiving the diagnosis result. Doctors can use backward chaining to recheck diagnosis results and provide feedback and advice to patients via email or online chat. Furthermore, by reviewing and comparing the patient's diagnostic result with the doctor's diagnosis comment/advice, the medical intern can learn diagnosis practice. Both diagnostic results can be saved in the PHR database for future use. In the event that online doctors are unavailable, the technology permits patients to seek diagnosis results straight from the system.

The Knowledge Base (KB) comprises data on disorders of ENT including COVID-19. The KB is designed as a MySQL DB accessed via php. These widely used tools ease maintenance tasks by DB administrators.

Users feel comfortable with interfaces that are easy to understand and help them perform their tasks efficiently. A task is an activity that is performed to reach a goal. The task analysis according to [8], was the basis for 1) gathering the requirements regarding the user interface, which is not a common procedure in the development of Knowledge Based Systems (KBS) [9], 2) developing, wire-framing and prototyping the site structure, and 3) performing usability testing with two appropriate user cohorts (medical personnel and potential patients). Usability requirements and appropriate testing has been neglected for a long time in the development of KBS because developers tend to focus on the knowledge basis. Overall, we identified four types of users: potential patients (ordinary users), non-specialist medical personnel (medical interns, doctors), specialist medical personnel (e.g., otolaryngologists), and administrators (e.g., Web admins).

The user interface was created in accordance with recommendations for the interface design of clinical decision support systems. It allows the identified user roles to employ COLDS within four layers based on the identified user roles:

1) patients: can access their personal EHR and communicate with doctors for rechecks, comments and further advice and can access COLDS through the question management interface,

2) specialists, e.g., otolaryngologists and ophthalmologists: can recheck patient's diagnosis result and provide comment and advice to the general practitioners, medical interns and patients via the chat room or email, can deliver the decision and recommendation of treatment to the patient and can analyze and evaluate medical intern's performance regarding anamnesis and diagnosis work, e.g., by applying the traditional Objective Structured Clinical Examination (OSCE) framework [10],

3) non-specialists, e.g., general practitioners and medical interns: can use patient's diagnosis result and doctor's comments as pieces of practical advice, and can review and recheck specialists' overriding diagnose results for further studies,

4) administrators: can change, add, or remove general information about disorders keeping it up-to-date and can access and review the log files to identify design issues.

In COLDS, we use well-established methods for acquiring patient data, such as asking specified and straightforward questions in Thai. Closed questions and two layers of language are used in the system: medical terms in the rule base and common terms in the user interface. After they have logged in to the system, users tick the chief complaint for further anamnesis, e.g., "sore throat" and "fever". By clicking on "Additional symptom", more specific complaints are shown, which are related to the body systems specified as the main complaints.

The system then uses the properties of disorders related to symptom details and prepares for the questioning task (yes/no) employing the question management process (Figure 2). The performance measures for COLDS were success rate, time to success and

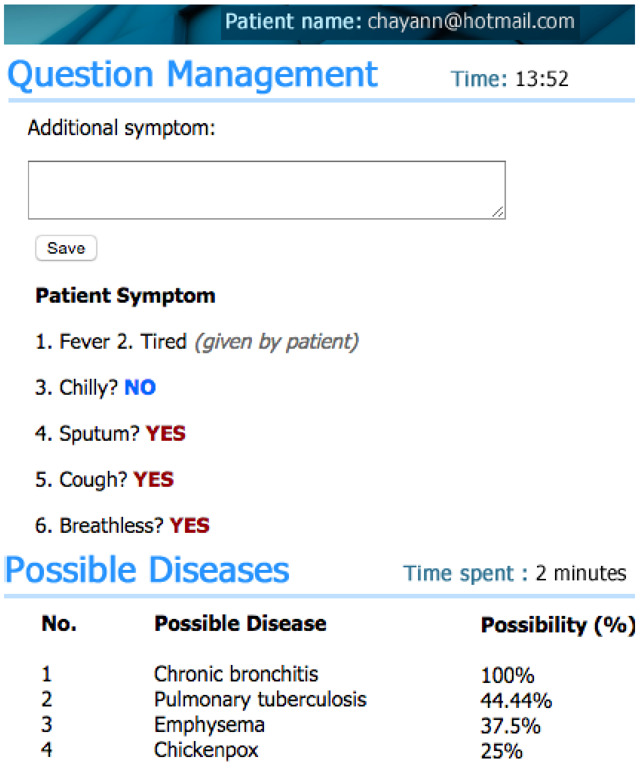

FIGURE 2. COLDS results (example)

number of errors. Additionally, we monitored self-reported satisfaction and comfort rating.

3. **Pilot Testing and Results.** The pilot test of COLDS was performed in October and December 2020 together with a Community Health Center in Sakon Nakhon Province, Northeastern Thailand. A second round of tests was performed in September and October 2021 after having made improvements to COLDS based on doctors' and interns' feedback.

Test objects (*and participants*) were 1) COLDS (*patients and medical staff*), 2) the user interface (*patients and medical staff*), and 3) the knowledge base (*medical staff*). Regarding the pilot testing, two different experiments were undertaken: continuous tests of the user interface over the development process, where twenty patients with varying skills regarding computer and Internet used the system to find defects in the steps of the development process, and system tests, which were performed with a group of fourteen medical interns and two doctors who are specialized in eyes, nose, throat and ears disorders, respectively. The tests were performed as follows.

1) The test of the CUI was performed as a continuous process accompanying software development and involved the twenty users to detect usability problems of the system. First, we divided the cohort into two groups: Group 1 with high computer knowledge group and Group 2 with lower background knowledge in computing. Group 1 was recruited to identify flaws and errors in the UI. They used it for patients and first checked at the design process, then at the prototype, and finally prior to the system test. Group 2 checked for readability of all texts, system accessibility, ease of use, response times, and comprehensible doctor input. Since the data related to this part of the test process are internal, we cannot report on them here. In the following, we focus on the results of the system test.

2) For the system test, medical interns were presented with the chief complaints (Table 1). To do the testing, medical interns had to think about one disease and tried to answer questions correctly according to the signs of the diseases. This resulted in that

twelve interns answered questions correctly, which can lead the system to a more effective diagnosis. Two professionals double-checked the system's results (the diagnoses) using a backward chaining procedure, which revealed how the system arrived at each diagnosis. The doctors not only double-checked the patients' answers to the questions, but also conducted a live anamnesis with the inpatients to determine the diagnosis in the traditional fashion.

4. **Diagnosis Performance Test.** Figure 3 shows an example of the question management process related to nose disorder diagnosis, which is fostered by a rule base. The rule base was revised by medical experts and can be used in practical tests by medical interns. The process is part of the medical question management for the automatic choice of appropriate questions for patients interface. For example, if the user/patient ticks "have fever" (1) the first question would start with "are you tired" (2), if yes goes to "anorexia?" (are you bored with food?) (3), if the answer is 'yes', then the question will be "feel chilly?" (4), and if all answers are yes for "sore throat?" (5), "pain in the ear/behind the ear?" (6), and "swollen jaw" (7), respectively, then the patient is diagnosed with mumps (disease A). However, if the answer in (2) is 'no', the question will move to another diagnosis question, which is "have sputum?" (2.1). If the answer is 'yes', the questioning carries on asking "breathless?" (2.2), and if the answers are 'yes' again asking for "cough?" (2.3) and "Muscle pain?" (2.4) respectively, then following these positive answers the patient is diagnosed with chronic bronchitis (disease B), but if the answer is 'no' for "cough?" (2.3), then the question moves to "cough with sputum only?" (2.3.1) and if the answer is 'yes', then the patient is diagnosed with emphysema (disease C). From 4, if the answer is 'no', the question moves to "high fever?" (4.1), and if all answers are 'yes' for "itchy?" (4.2), "rash?" (4.3) and "blisters?" (4.4), respectively, then the patient is diagnosed with chickenpox (disease D). From sore throat (5), if the answer is 'no', then the question moves to "have sputum?" (5.1), and if all answers are 'yes' for "cough?" (5.2),

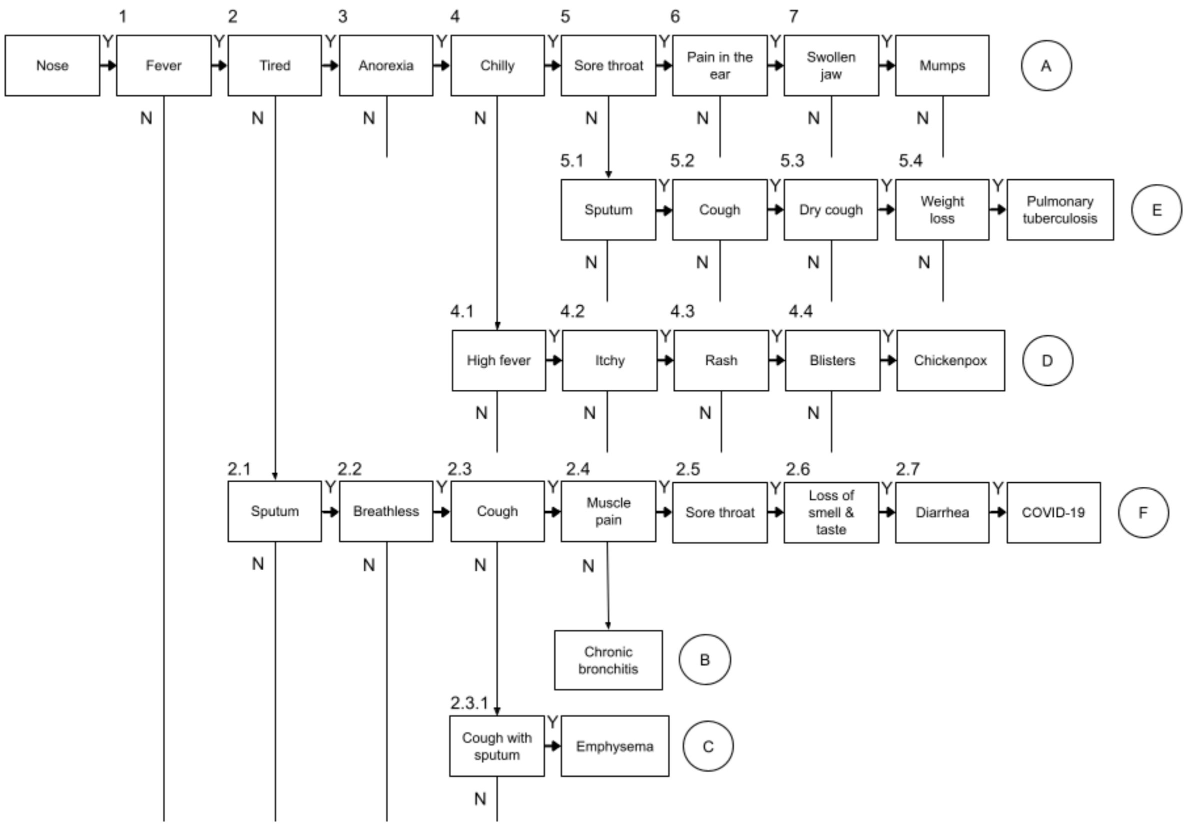

FIGURE 3. Examples of the COLDS question management process

"dry cough?" (5.3), and "weight loss?" (5.4), respectively, then the patient is diagnosed with pulmonary tuberculosis (disease E). Similarly, conditions are analyzed for possible COVID-19 symptoms (F). With the help of experts, we have constructed a number of rule classes for eye, nose, throat and ear disorders, and these have been revised by doctors with expertise in these medical areas.

However, if the diagnose steps did not work as in Figure 3, another example is provided referring to anorexia (3): if the response is "no", then the system takes all the previous answers to conditions as "yes" and tries to look for other disease diagnosis paths related to the maximum number of conditions matched and then asks questions until the result has been found. To perform this analysis, the system reduces the "yes" conditions one by one and looks for other possible diagnosis paths. For example, if "fever" and "tired" conditions together were not in any of the other diseases groups (conflict case), then the system cuts off the "tired" and uses the "fever" condition and looks for a group of diseases that is related to fever and chooses the second highest frequency of relating conditions. The question management process starts again until the disease has been found. Otherwise, the message of "You are advised to present your complaints to a doctor" will be shown together with a location aware suggestion of a nearby hospital or clinic, if available.

The findings demonstrate an 87 percent consistency/efficiency between the system's and doctors' diagnoses. Errors result from unknown diseases for which the symptoms provided did not match any cases, owing to the fact that disease rule bases were missing and did not connect to any of the diseases that were available. Patients, doctors, and interns also evaluated the user interface. Patients received a score of 4.29/4.35, interns a score of 4.53/4.63, and doctors a score of 4.33/4.55 for ease of use/overall satisfaction. On a scale of 1 to 5, where 1 indicates high discontent and 5 indicates high satisfaction, the results suggest that 4.38 for ease of use and 4.51 for satisfaction (with 95 percent confidence interval: 4.51-4.59).

Table 1 shows the evaluation results of the performance testing. The number of successful results states how many chief complaints could be diagnosed by COLDS. After specialists rechecked the results, they accepted or dismissed results. The percentage gives the ratio of accepted to successful diagnoses.

TABLE 1. Chief complaints and related diagnose success from COLDS

Chief complaint	Number	Number of successful diagnoses	Number of successful diagnoses accepted by specialists	Percentage of accepted diagnoses
Ear	6	6	5	83.3%
Nose	9	8	7	87.5%
Throat	6	6	6	100.0%
Eyes	4	3	2	66.7%
Sum	25	23	20	87.0%

According to Kirkpatrick [11], testing of training efforts can be evaluated at four outcome levels: 1) reactions, 2) learning and attitudes, 3) behavior, and 4) organizational results. In the context of this research, we focused on investigating reactions, learning and behavior of the participants, i.e., medical staff and patients. The usability test was based on a sound procedure that has been outlined in [8]. Here, we can only give a summarizing report on the test plan and its subsequent outcomes, which are as follows.

Objectives. COLDS usability testing aimed at assessing human performance on key tasks in terms of effectiveness and efficiency (especially for frequently performed tasks) and at finding critical errors affecting patients' well-being as well as monitoring user satisfaction with the system performance. Potentially critical conditions stated by patients are flagged in the knowledge base and are subject of evaluation in the diagnose evaluation process.

Methodology. The participants of the usability study were grouped by their experience with the Internet. Related to the potential study participants, we derived some insight from online usability experts who have suggested, among others, the following practices [6]: 1) for insights regarding computer and Internet experience, ask questions about the behavior, no self-assessment, 2) ask precise questions, and 3) use a pilot test for the screener.

This methodology was applied to the core user groups: potential patients (i.e., ordinary users), non-specialist medical personnel and specialist medical personnel. The outcomes were measured as success rate, time to success, and number of errors. Additionally, we analyzed self-reported satisfaction, ease of finding information and ease of use ratings for all groups including the medical staff.

Materials. The key use scenarios in the tests were 1) in-lab testing by participants (observe and record) [12], and 2) online testing by participants (review). We used data from EHR made accessible on COLDS for both the medical staff (non-specialists and specialists) and the potential patients. The EHR data comprised general data (name, address, gender, age, general conditions) as well as past data on ENT, eyes and other complaints.

The patient answers the questions provided by the system and also adds more detail of signs if necessary until appropriate diagnosis results can be derived with sufficient statistical values (Figure 2). These values are calculated by number of signs found and given by user answers divided by the number of all signs of disorders found. In Figure 3, where the patient answered 'yes' to the following questions: have sputum? Cough? And breathless? Consequently, the system gave as possible values: Chronic bronchitis (100%), Pulmonary tuberculosis (44.44%), Emphysema (37.5%) and chickenpox (25%). The threshold for results being shown to patients is 50%, whereas doctors and non-specialists can see all results. The user/patient can seek diagnosis results directly without waiting for the doctor's online recommendation/advice in some basic or typical diagnosis scenarios (mostly trivial concerns). In the instance of more serious problems, the doctor would review the patient's symptoms, analyze the diagnosis, and then either provide treatment and advice to the patient or ask the patient to present their situation to a medical specialist.

The user interface of COLDS allows users to update their personal data in the EHR, e.g., their marital status, address, phone numbers, and occupation. Moreover, the system is connected to the EHR via the Patient ID and provides the past medical history of diagnosis/analysis, e.g., allergies, childhood illnesses, immunization, surgeries, blood transfusions, severe and chronic diseases, and accidents. Additionally, the complete lab data for the patients are available to the system, including pulse rate, body temperature, blood pressure, weight (with current BMI derived), and can be used by the system and doctors for diagnosis.

Another interesting metric for assessing the merits of COLDS is the time it takes each group of users to get to the final diagnosis. The data are as provided in Table 2.

Table 2. Time consumption using COLDS

User group	Average time spent	Maximum time spent per session	Minimum time spent per session
Patients	4:47 min	8:17 min	1:13 min
Non-specialists (medical interns)	3:37 min	4:41 min	2:22 min
Specialists	4:44 min	5:29 min	3:58 min

5. **Discussion.** By using COLDS, the time burden for successful automated diagnoses is mainly on the individual patient's side, and the doctors are affected on a much lower level (Table 2). This is usually welcomed by both specialists and non-specialists. We do not present data on time spent for unsuccessful automated diagnoses because those need specific interpretation as to why and how the non-specialist or the specialist came to their diagnosis with the help of EHR data (or not, in which case patients had been sent to clinics). On the other hand, patients do not need to travel to the clinic or local community health center, which would have been cumbersome for many. Without using systems as colds, patients may seek help from nearby 'village doctors' with limited knowledge and inadequate tools, materials, and skills. In case of medical issues, the time required to use colds may not be a significant barrier.

New sociotechnical systems cause concern if not fear among people involved, especially in case of automated systems, which are often seen as job killers. It was, therefore, necessary to promote the goals and potential achievements of the COLDS project as publicly and openly as possible in advance to promote its application even at the test stage.

6. **Conclusions and Further Work.** The results of design-and-create research for a Web-based telemedicine diagnosis system with a Conversational User Interface as the frontend and an expert system are provided in this paper. The technology has been carefully designed to save time during the anamnesis procedure for users/patients with ENT and ocular disorders, as well as clinicians. Users/patients can utilize the system as an outgoing diagnosis system from their homes, or as an inbound diagnosis system to help allocate a patient presenting to a specialist in the medical facility. The user interface has been designed as simple as possible, with the result that after three rounds of thorough usability testing, users/patients reported no faults. The system was tested in two ways: 1) usability testing with diagnosis results for medical interns and doctors, and 2) interface testing for all types of users, including those with and without computer capabilities.

REFERENCES

[1] S. Ausayakhun, A. H. Skalet, C. Jirawison, S. Alisayakhun, J. D. Keenan, C. Khouri, K. Nguyen, P. S. Kalyani, D. Heiden, G. N. Holland and T. P. Margolis, Accuracy and reliability of telemedicine for diagnosis of cytomegalovirus retinitis, *American Journal of Ophthalmology*, vol.152, no.6, pp.1053-1058, 2011.

[2] K. Denecke, M. Tschanz, T. L. Dorner and R. May, Intelligent conversational agents in healthcare: Hype or hope?, in *Healthcare of the Future*, T. Bürkle et al. (eds.), IOS Press, 2019.

[3] K. Denecke and J. Warren, How to evaluate health applications with conversational user interface?, *Digital Personlized Health and Medicine*, vol.270, pp.976-980, 2020.

[4] G. Falkman, M. Gustafsson, M. Jontell and O. Torgersson, SOMWeb: A semantic web-based system for supporting collaboration of distributed medical communities of practice, *Journal of Medical Internet Research*, vol.10, no.3, DOI: 10.2196/jmir.1059, 2008.

[5] S. J. Swierenga, Incorporating Web accessibility into the design process, in *The User-Centered Design Casebook*, C. Righi and J. James (eds.), Morgan-Kaufmann, 2007.

[6] W. Quesenbery, Dimensions of usability, in *Content and Complexity: Information Design in Technical Communication*, M. Albers and B. Mazur (eds.), Erlbaum, 2003.

[7] J. Horsky, G. D. Schiff, D. Johnston, L. Mercincavage, D. Bell and B. Middleton, Interface design principles for usable decision support: A targeted review of best practices for clinical prescribing interventions, *Journal of Biomedical Informatics*, vol.45, pp.1202-1216, 2012.

[8] J. A. Landay, N. Oliver and J. Song, Conversational user interfaces and interactions, *IEEE Pervasive Comput.*, vol.18, no.2, pp.8-9, 2019.

[9] M. Freiberg, A. Striffler and F. Puppe, Extensible prototyping for pragmatic engineering of knowledge-based systems, *Expert Systems with Applications*, vol.39, pp.10177-10190, 2012.

[10] D. D. Nulty, M. A. Mitchell, C. A. Jeffrey, A. Henderson and M. Groves, Best practice guidelines for use of OSCEs: Maximising value for student learning, *Nurse Education Today*, vol.31, no.2, pp.145-151, 2011.

[11] D. L. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, Berrett-Koehler, 1998.
[12] A. C. Li, J. L. Kannry, A. Kushniruk, D. Chrimes, T. G. McGinn, D. Edonyabo and D. M. Mann, Integrating usability testing and think-aloud protocol analysis with "near-live" clinical simulations in evaluating clinical decision support, *International Journal of Medical Informatics*, vol.81, no.11, pp.761-772, 2012.