

Improving clinical practice using a computerized  
clinical decision support system for diagnosing rare  
diseases: literature review, challenges, and possible  
paths forward

Radu Drăgușin

Paula Petcu

October 2010

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Background</b>	<b>3</b>
2.1	Diagnostic process . . . . .	3
2.2	Medical information resources used by clinicians . . . . .	3
2.3	Computerized clinical diagnosis decision support systems . . . . .	5
2.4	Rare diseases . . . . .	5
<b>3</b>	<b>State of the art in medical diagnostic systems</b>	<b>7</b>
3.1	History of clinical decision support systems . . . . .	7
3.2	Types of clinical decision support systems . . . . .	8
3.3	Available clinical diagnostic decision support systems . . . . .	8
3.4	Rare diseases and CDSSs . . . . .	10
<b>4</b>	<b>Review of guidelines for implementing and evaluating CDSS</b>	<b>12</b>
4.1	An ideal diagnostic support system . . . . .	12
4.2	Evaluation of clinical software . . . . .	13
4.3	Social implications . . . . .	14
4.4	Users' perspectives . . . . .	15
<b>5</b>	<b>Future work</b>	<b>16</b>
5.1	Support decision system for diagnosing rare diseases . . . . .	16
5.2	The development cycle . . . . .	16
5.3	Sources of information for the rare disease project . . . . .	17
5.3.1	Text-based information sources . . . . .	17
5.3.2	Non-text-based information sources . . . . .	18
5.4	System design . . . . .	19
5.4.1	Querying interface . . . . .	19
5.4.2	Information retrieval and processing . . . . .	19
5.4.3	Evidence strength . . . . .	20
<b>6</b>	<b>Conclusion</b>	<b>21</b>

# Chapter 1

## Introduction

Diagnosing rare diseases can be a difficult and time consuming task for a physician. When encountering the symptoms of a rare disease, the physician usually has little or no prior experience with similar cases. Although time is of major importance in most of the cases of rare diseases, the patient is usually referred to a specialist. The lack of the correct diagnosis or the delay caused by going from one specialist to another can lead to undesirable outcomes. A physician could query the symptoms on a specialized database such as PubMed, however, this general databases return numerous unnecessary results. Thus, it could prove to be useful for the physician to have a tool specialized in matching symptoms with rare diseases.

The objective of this report is to review the state of the art in clinical decision support systems (CDSS) and identify guidelines for designing, implementing and evaluating such a system, with a focus on rare diseases. Based on these findings, we present possible directions for future development of a CDSS for diagnosing rare diseases.

The report is organised as follows: Chapter 2 describes the various processes underlying the research for developing a clinical diagnosis support system for diagnosing rare diseases. Chapter 3 outlines typical definitions of CDSSs and reviews the state of the art of these, including their in relation with rare diseases. Finally, Chapter 4 provides guidelines for implementing and evaluating a CDSS, Chapter 5 proposes some future directions of research and Chapter 6 concludes the review.

## Chapter 2

# Background

### 2.1 Diagnostic process

In order to design a tool to improve the accuracy of medical diagnosis, it is first of all necessary to understand the cognitive processes underlying it [1].

To make medical decisions, the physician must combine the patient's data with two types of medical knowledge: a low-level knowledge about the structure and the function of the body, about diseases, their causes and treatments; and a second high-level knowledge gathered from clinical experience [2].

Given the patient's symptoms, the clinician uses his medical knowledge to formulate hypotheses about the cause. Based on their likelihood and perhaps treatability, one or two hypotheses are selected and further information is collected in order to test them. With the new acquired data, a hypothesis may be confirmed or other hypotheses may be further explored. The process ends in a decision to intervene [2].

The diagnostic process can be viewed as a Bayesian inference where the prior probability of a condition is revised in accordance with new evidence [3].

It is the second type of knowledge (clinical experience) that enables clinicians to organize the evidence and make 'wise' decisions [2]. However, even experienced clinicians sometimes need guidance in remembering, confirming or acquiring medical knowledge [2]. In such situations, the clinicians may turn to several resources of medical information.

### 2.2 Medical information resources used by clinicians

Today, clinicians are overwhelmed by the great deal of medical information available. A typical physician is estimated to use approximately 2 million pieces of information when taking medical decisions [2, 4], and the amount of

medical knowledge is rapidly increasing<sup>1</sup>. It was also reported that clinicians have at least one medical knowledge-related question per patient, and that more than half of these remain unanswered [4, 5]. Knowing that a leading cause for medical errors are delayed or erroneous decisions [6], it is important to understand the reasons clinicians do not pursue the answers. The main reasons for not pursuing an answer seem to be time constraints, lack of adequate resources [5], difficulty formulating a question, difficulty selecting an optimal search strategy, failure of resources to cover a topic, inadequate synthesis of multiple resources, and uncertainty about when all relevant information was found and the search can stop [7].

Regarding the gap between evidence and practice, it has been studied that there is often an unacceptable delay between the confirmation of a medical research finding and its use in clinical practice [8]. It was also found that it takes an average of around five years for the published guidelines to be adopted in routine practice [9].

To overcome these problems, the clinicians use information resources to supplement their knowledge and clinical experience, and to keep up to date. They have traditionally used journals, textbooks, and discussions with colleagues to find information, but more popular have now become the electronic online resources. Data reported in several studies reveal that most physicians (results varying between 72% and 95%) regularly use the Internet to identify medical information [10, 11], and in one of these surveys, 51% of the participating physicians declared that the Internet influence their healthcare decision [11].

Many medical articles are now available online on the websites maintained by journals, although there are some that require registration or charge for viewing the full material. However, many of the articles are also available through non-journal web sites and online repositories maintained by the author or author's institution [12]. Some medical professionals have access to full-text journal articles through institutional subscriptions.

MEDLINE is the authoritative repository of medical bibliographic information containing basic article-related information such as title, authors, publication information (name, date and type), abstract, language, metadata and other data. The MEDLINE database is accessible through search interfaces. The most popular of these is PubMed, which allows the use of boolean operators, search templates and other complex querying techniques.[13]

## **Review of existing evaluations of different sources the clinician use when searching for medical information**

A 2004 survey reported that most physicians (92%) favour using a specific website over commercial search engines (such as Google) for gathering med-

---

<sup>1</sup>Between 2,000 and 4,000 MEDLINE citations are added monthly, with 712,000 new references added in 2009 alone. MEDLINE Fact Sheet

ical information. Nearly one third (32.3%) of the physicians participating in the survey reported that they use edited data sources (such as UpToDate, Medscape, WebMD, MD Consult, and eMedicine), more than one quarter (27.3%) favor research databases (such as PubMed) which provide access to medical journal publications, and a small percentage use sites specialized to their area of expertise (2.9%) or use medical web site portals (3.1%) [14].

In a 2009 study of healthcare professionals from Finland, it was discovered that the majority of respondents used electronic databases in their practice and that more than 80% of them used more than four such databases. Among the most popular for physicians were Google (82%) and PubMed (74%). However, 93% of the nurses used Google, but only 42% used PubMed [15]. This indicates that PubMed is especially useful for highly qualified healthcare professionals.

It was observed that Google is excellent at finding resources where symptoms and signs co-occur and that humans are efficient in filtering relevant documents[16]. However, problems arise either in the case of complex diseases that exhibit no specific symptoms or common diseases with an unusual case of manifestation, or when the clinician is flooded with non-relevant results [16].

### **2.3 Computerized clinical diagnosis decision support systems (CDDSS)**

A computerized clinical decision support system (CDSS) is an information system designed to improve clinical decision making. There are several types of CDSSs, but this study focuses on the diagnostic systems, which assist the clinician with the process of differential diagnosis. These systems are intended to remind the clinician of diagnoses that they might not have considered, or trigger their thinking about related diagnostic possibilities.[17]

These type of systems usually take as input patient information such as signs, symptoms, past medical history, laboratory results values and demographic characteristics, and generate a list of possible diagnoses, ranked probabilistically. CDDSSs do not attempt to replace the clinician by providing one correct answer, but they work by interacting with the clinician in selecting a set of plausible alternative diagnoses.

### **2.4 Rare diseases**

Rare diseases are generally defined by the number of people affected by the disease, but other definitions include factors such as the severity of the disease or the existence of a treatment. To be classified as a rare disease, the prevalence of that disease has to be very low, but different countries select different thresholds to establish this division. A disease that may be

considered common in one part of the world may be labelled as rare in other parts.

In Europe, a disease is defined as rare when it affects fewer than 5 persons in 10,000, i.e. less than one person in 2,000. In Denmark, a disease or disability is considered rare when there are less than 1,000 people affected nationwide i.e. approximately one in 5,000 inhabitants. In Norway, a condition is rare when there are up to 100 people affected by the disease in one million inhabitants, i.e. one in 10,000. The Swedish National Board of Health and Welfare has a database of rare diseases defined as affecting "fewer than 100 people per million, and which lead to a marked degree of disability."<sup>2</sup>

In United States, a rare disease is defined<sup>3</sup> as "those which affect small patient populations, typically populations smaller than 200,000 individuals in the United States" – that is, around 1 in 1,500 people.

According to the National Institutes of Health Office of Rare Diseases, there are close to 7,000 known rare diseases in the world and about 25-30 million people in the U.S. are affected by one<sup>4</sup>. The European Organisation for Rare Diseases, EURORDIS, estimates that there are between 6,000 and 8,000 rare diseases, and that around 30 million European Union citizens may be affected by a rare disease, most of them (around 80%) of genetic origins<sup>5</sup>. It is estimated that around five new rare diseases are described each week in the medical literature.<sup>6</sup>

---

<sup>2</sup>Rarelink.no is a Nordic site that contains links to information on rare disease that have been published in Norway, Sweden and Denmark About rarelink.no

<sup>3</sup>According to the Rare Disease Act of 2002. Full text

<sup>4</sup>The Office of Rare Diseases Research, National Institutes of Health. About ORDR

<sup>5</sup>EURORDIS - "The voice of Rare Disease Patients in Europe" - is a non-governmental patient-driven alliance of patient organisations and individuals active in the field of rare diseases EURORDIS website

<sup>6</sup>Orpha.net - "The portal for rare diseases and orphan drugs" About rare diseases

## Chapter 3

# State of the art in medical diagnostic systems

### 3.1 History of clinical decision support systems

It has been repeatedly asserted in the literature of computerized medical diagnosis that clinical decision support systems have a great potential to reduce medical diagnostic errors and improve quality of care [1, 18, 19, 20, 21]. Despite the existing results of improved medical care, their demonstrated utility and various investments in the field, health care institutions have been slow to incorporate these systems in the work environment[22].

The first attempts to create software programs that aid in diagnosing medical conditions started more than five decades ago [23, 24]. The first generation of CDSSs (such as MYCIN, QMR, Iliad or DXplain) used pre-compiled knowledge bases of symptoms, signs and laboratory findings for a list of syndromes and diseases, and the user entered his patient findings through a menu of choices. The program then used Bayesian logic or pattern-matching techniques to suggest possible diagnoses [25]. Although these systems proved to be helpful to the clinician in experimental settings, these early systems did not gain widespread acceptance in clinical use. The difficulties encountered were presumably the considerable amount of time needed to input the clinical data and the system's limited sensitivity and specificity [25].

Many experimental programs have been developed over the years, but only few of them were integrated in the clinical environment. It was observed that there is a pattern in the development cycle of many of these systems; the new models were tested and their characteristics were of interest, but then they were abandoned when developers found out that further development was much more difficult than modelling the prototype.[23]

Significant research progress has been made since the idea of the computer-based CDSS emerged, however, several barriers continue to interfere with



their effective implementation in clinical settings [24]. A 2005 systematic review concluded that CDSSs improve practitioner performance but there is insufficient evidence to determine the effects on patient outcomes [26, 24].

### 3.2 Types of clinical decision support systems

Computerized algorithms can be used to provide decision support for a variety of clinical conditions. Clinical decision support systems are developed to support medical functions (or tasks) such as alerting (e.g. highlighting abnormal values), reminding (e.g. to schedule a surgery), critiquing (e.g. reviewing a prescription), interpreting (e.g. electrocardiogram interpretation), predicting (e.g. risk of mortality), diagnosing (e.g. producing a differential diagnosis), assisting (e.g. in selection of antibiotics), and suggesting (e.g. generating suggestions for adjusting medical equipment).[4]

Clinical decision support systems can also be described along several other axes, besides their specific tasks. The systems can differ in the timing of providing support (before, during, or after the clinical decision is made), scope of the system (general or targeting a specialty), setting (ambulatory or inpatient care), access (integrated with an electronic health record system or stand-alone systems). The systems can also be described in terms of implementation, as being knowledge-based or non-knowledge-based systems (employing machine learning or other statistical pattern recognition approaches [27]), and in terms of the output, as active or passive system (e.g. actively providing alerts or reminders, or passively responding to user input).[27, 28]

### 3.3 Available clinical diagnostic decision support systems (CDDSS)

The development, evaluation and application of the systems described in what follows have been extensively presented in literature [17]. Regardless of implementation, CDDSSs usually take as input clinical signs, symptoms, laboratory results or demographic characteristics, and output diagnostic recommendations [29]. The systems differ in the data used to determine their probabilities estimates, the diseases and clinical findings covered in their knowledge base, the representation they use for describing data, and the algorithms they use to process and analyse the data [17]. Most knowledge-based systems are comprised of three parts: the knowledge base, the inference engine, and the user interface [27]. Systems that do not use a knowledge base use machine learning, recognizing patterns in data [27].

Examples of diagnostic decision support systems are: INTERNIST-I and MYCIN (two of the first expert systems to be developed in the 1970s, rule-based systems that follow "if-then" rules), DXplain (first version released

in 1986), Problem Knowledge Coupler (PKC), Iliad, Isabel, QMR (Quick Medical Reference), Global Infectious Diseases and Epidemiology Network (GIDEON), SAPHIRE and MedQA [30, 25, 29, 31]. Miller observed in the 1990's a trend in the clinical decision support systems of becoming more specialized and focused, integrated in the clinical workflow [32]. He also observed that the importance of evaluating these systems has raised and that developers begin to include patient preferences and evidence-based knowledge representation in their design [32].

In what follows, we briefly describe some of the most popular of the systems currently available. Giving a detailed description of these systems is not the focus of our report as they are exhaustively described in previous literature.

Isabel is a web-based CDSS that accepts both free-text entry and key findings based input. It uses a thesaurus for recognition of terms, uses natural language processing strategies and pattern-matching algorithms to compare findings with a selected reference library (comprised of 6 medical textbooks and 46 major journals of general and specialty medicine) and produces a total of 30 suggested diagnoses (with links to the authoritative texts).[25]

Iliad and QMR are two mature systems that were extensively covered in the literature and are available commercially for use by physicians. Users of Iliad can enter clinical findings about a patient using an interface that allows both free-text and menu-based input. Estimated probabilities are associated with the diseases generated by the system. The representations of each disease can also be browsed. QMR also offers functions such as comparing and contrasting pairs of diseases. In QMR, the relationship between findings and diseases is represented on a 5-point scale.[33] QMRs predecessor is INTERNIST-1 [30].

MedQA is a "medical definitional question answering system" that generates short (paragraph-level) answers to definitional questions (i.e. questions of the format "What is X?")[34, 31].

Denekamp et al [30] present a comprehensive description and categorization of the diagnostic systems currently available. They break down the characteristics of the CDSS in several important features (e.g. the systems' sources of knowledge, computational model, explanations provided), including the consideration of all hypotheses, where DXplain has an advantage: the rare diseases are displayed separately from the rest of the diseases. DXplain is a mature and evolving system; the development of DXplain began in 1984, its distribution in 1987, and in 1996 it was switched to the online version [35]. A valuable insight from the creators of DXplain is that, although they implemented numerous advanced features in the later versions, most of these features were used by an extremely small percentage of users [35].

### 3.4 Rare diseases and CDSSs

In 1987, in a review of diagnostic logic, Macartney stated that "probably the most suitable [computer diagnosis] application is in the diagnosis of rare syndromes, where there is a real problem of human memory and the collation of small snippets of information from diverse sources. Cases can be added from reports in journals as well as from the experience of collaborating centres, thus pooling information that could never effectively be accommodated in the memory of a single clinician"[36].

A system for the computerized storage and retrieval of information on rare dysmorphic syndromes was developed in early 1980's. The London Dysmorphology Database described in the 1984 paper was developed with the aim of providing the clinician with a manageable list of possible genetic syndromes for a particular case, with references, and the possibility to register undiagnosed or unreported cases [37]. An online and updated version of the project was published in 1999 <sup>1</sup>.

Thus, it seems reasonable to say that the literature supports the idea that a computerized medical diagnosis system is well suited for diagnosing rare diseases. The average clinician would not be able to remember or recognize the majority of the rare diseases, and even an expert clinician would find it difficult to recall all of them.[37] A diagnosis decision support system could help the clinician identify the most plausible diagnosis, provided that the system is able to produce an evidence-based list of the most significant rare diseases matching the patient's symptoms.

#### Using the Google Search Engine as a diagnosis tool for rare diseases

There have been several cases in which patients have correctly diagnosed themselves or close relatives with a rare disease. One such report is presented in [38], where in two different cases, parents concerned by the lack of diagnosis searched Google for symptoms and discovered a rare disease that they suspected to affect their children. When informed of this suspicion, doctors ordered the appropriate tests which confirmed that indeed the children suffered from a rare disease for which early treatment was essential.

Nevertheless, despite these success stories, there have been cases of harm associated with the use of health information on the Internet. Section 4.3 of this report discusses the social implications of the widespread availability of health information on the Internet.

There are a couple of information search engines that have been devel-

---

<sup>1</sup>The online database contains information on over 4450 syndromes and has an integrated photo library that shows the main dysmorphic features of the syndromes (accessed October 2010) <http://www.lmdatabases.com/>

oped for the sole purpose of searching information on rare diseases<sup>2</sup>. These perform the searches only through a few major rare disease information sources and databases, thus eliminating irrelevant results that would sometimes appear in a Google search and that would make finding the answer difficult.

---

<sup>2</sup>The Rare Disease Search Engine <http://www.raredisease.org/> and the Rare Disease Communities Search Engine <http://www.rarediseasecommunities.org/en/search>

## Chapter 4

# Review of guidelines for implementing and evaluating a CDSS

### 4.1 An ideal diagnostic support system

The literature on computerized medical diagnostics provides various helpful general recommendations on how to develop a successful CDSS. Many of these recommendations are produced as a result of a survey or a focus group, or from lessons learned by the developers of such systems. Based on these, in what follows, we have outlined the factors that may influence the success of a CDSS.

*A system should perform tasks that a doctor could not* [36]. Mimicking the physician is not enough; the systems may have a greater acceptability potential if they prove to perform better in a specific task than the average physician. The early systems developed for improving the medical diagnosis were expert systems build on a model that tried to simulate human thinking. This model is however not enough for the CDSS to succeed [27]. The systems must provide additional, useful information about the patient [3].

*Knowledge in the system must be current, accurate and verifiable* [39]. The system should be transparent to the user, offering meaningful evidence at each step, and should be continuously updated with the latest research evidence [27].

*Knowledge in the system must be easily accessible* [39]. The system should be available at the time and place of care, where the doctors see patients and decisions are made [27], and should not require a lengthy log-on or multiple question-answering steps [9].

*Monitor impact, gather feedback, and respond with improved software* [9]. Also monitor proper utilization of the installed clinical decision support system, and assure that the users are properly trained [27].

*Make sure that the users understand the limitations of the CDSS* [27]. The clinicians should be informed from the start of both the strengths and limitations of the product. On the other hand, the developers should cover the situation in which the system does not contain a piece of information (a disease) that matches the findings (patient's symptoms).

*The applications should anticipate clinician needs* [9]. Assist the clinician in the process of diagnosis by anticipating his needs (for example, by suggesting new symptoms matching the results), and by delivering the needed information in real time.

*Speed is important.* Studies found that physicians would search for up to two minutes for an answer, giving up searching if an answer was not found after two minutes [11, 34]. One study evaluating the use of PubMed by clinicians showed that it takes approximately half an hour to answer a clinical question using PubMed [34]. This suggests that PubMed searching can not be used in the clinical setting, but only after clinical hours [34].

Other challenges that are important to be met in the next generation of diagnosis decision support systems are the creation of Internet-accessible repositories, the use of free-text information, the prioritization and filtering of recommendations to the user, summarization of patient information, integration with electronic health record (EHR) systems, and the mining of large clinical databases [27, 20].

The use of some types of clinical decision support systems has increased over the years (reminders or alerts). However, the use of diagnostic decision support systems has been less popular [3, 8] because clinicians do not show a strong need for such a system, because it is harder to implement than the other types of CDSSs, and because the current systems would produce many irrelevant results [8]. In order for the systems to be useful, it was suggested that they should perform well, especially for difficult cases, and be flexible on the input [23].

## 4.2 Evaluation of clinical software

The evaluation of a clinical decision support system can be seen both a quality control mechanism, and as feedback for a new development cycle [32]. The first evaluation step that should be considered is the identification of the exact decision problem which the system should resolve (a prototype may be necessary before a clear definition can be articulated) [19].

CDSSs should be tested throughout their whole life cycle, starting even before they are implemented [18]. It was proposed that the evaluation of medical systems should be conducted in two stages: laboratory and field testing. The 'laboratory testing' phase is necessary in order to show that the system is safe and that patients can benefit from its usage. The 'field testing' phase is necessary in order to show if they have real value to patients

and clinicians [19].

It was also suggested that the evaluation of the CDSS should be made by considering three perspectives of the software - structure, process and outcome - and that there are two protagonists that must be considered for the evaluation process - the prospective user and the expert authority. In this suggested framework, the user would typically be a nurse or junior doctor, while the expert would be "the authority who sanctions the use of the system", a consultant. [19]

Following the evaluation model described above, for each type of protagonist (user and expert) three categories of laboratory testing questions should be asked: about the structure, process and outcome. Similarly, for field evaluation Appendix A provides examples of such questions.

### 4.3 Social implications

A growing number of researchers suggest that the successful adoption of a computerized decision support system does not depend only on the technology, but also on the social, political, organizational, and practice-related factors [40]. It was proposed that elucidating these contextual factors could facilitate the successful adoption of such systems in the clinical setting and result in a better system design [40].

Regarding the introduction of the computer in the interaction between patient and clinician, many [41, 42] feared that this will disrupt the personal setting in which the encounter between the patient and clinician takes place. However, studies [22, 21] have shown that the majority of the patients agree with the clinician using a computer during the consultation, and many of the patients perform Internet searches to look for health-related information themselves.

Patients use the web (and subsequently, the Google search engine [16]) to look for more information on their disease or symptoms, and in some cases, even try to diagnose themselves. However, in general, patients see the web just as an additional source of health information, and not a replacement for their doctor [43]. Thus, to avoid undesired situations, the doctor could then guide the patient to the correct medical information on the Internet, rather than ignoring the fact that the user will most probably search the Internet for more information. This results in patients assuming a more active role in the healthcare decision process.

An increase in patient use of medical information has been recorded in recent years [44]. Although there have been successful cases in which patients have correctly diagnosed themselves using the information available on the Internet [38], there have also been cases of harm produced by use of inaccurate information or from misunderstandings [45]. In these cases, the patient can suffer physically, emotionally, or financially [45]. These conse-

quences are especially important when the patient is wrongly self-diagnosed with a rare disease <sup>1</sup>.

Another concern, introduced by clinicians, was that the CDSS would increase consultation times [22]. In a study from 2006, the majority of the participating clinicians admitted that they were behind schedule most of the time by more than 20 minutes [46].

A further concern was that the system could provide advice in the form of an appealing but incorrect diagnosis, and if accepted by clinicians, the system's effect may be damaging [33].

On the other hand, a report has enumerated the several benefits of the interaction between the clinician and the information system: the interaction occurs when clinicians learn new information about patients and when clinical decisions are made, the information needs may be predictable and resolved, sometimes even automatically [6].

#### 4.4 Users' perspectives

It has been well described in literature that in order to design a successful CDSS it is important to first consider the users' needs, their attitudes and preferences, and their information needs [47, 42]. User resistance seems to be an important obstacle in the adoption of CDSSs in the clinical setting [42].

A study with clinicians published in 2010 showed that lack of time was perceived by the participants as the biggest barrier to using a CDSS [48]. This finding has been shown in previous studies as well [49], and thus it should be an important consideration when designing a CDSS. Clinicians reported a desire for the best evidence, supported by direct links to relevant and popular papers, such that there is a balance between relevance, popularity and quality [50].

Another study, concerning clinicians in Finland, reported enthusiasm for the use of CDSSs in practice, but also uncovered worries for the usability and the possible implications in the doctor-patient interaction [42].

One concern reported by some clinicians was that the use of a CDSS could result in more diagnostic tests being ordered. On the other hand, some clinicians acknowledged the usefulness of a CDSS in difficult cases [42].

---

<sup>1</sup>Cyberchondria New York Times article



## Chapter 5

# Future work

### 5.1 Support decision system for diagnosing rare diseases using vector space model and medical text mining

The idea of creating a system specifically for supporting diagnosing rare diseases was probably first described in *Support decision system for diagnosing rare diseases using vector space model and medical text mining* [51]. The system developed by M. Andersen and H.G. Jensen uses a rare diseases database containing information gathered from a rare diseases website, and has the ability to present the physician with a list of top 20 possible diagnoses given a list of symptoms. They used the bibliographic database MEDLINE together with PubMed, the public gateway for accessing this database, for accessing the article-related information on which the outcomes are based.

Their tests show that in around 60% of the cases, the system places the correct rare disease name in the top 20 results. However, their evaluation of the system could have benefited from a comparison with other systems or techniques (such as searching Google) and from a stronger statistical evaluation, ensuring that their assessment of performance is statistically significant. Moreover, it was shown that CDSSs evaluated by their developers have a tendency to show better results, which might indicate a bias [52, 53].

### 5.2 The development cycle

The development of the system should follow an iterative build-test-refine cycle [19]. A six-step process that can be used in implementing the system would begin with (1) identifying stakeholder goals; (2) identify available clinical systems; (3) select CDSS interventions; (4) build them; (5) deploy them; and (6) analyse their effect [8]. Again, the feedback collected from step 6 will help refine the system [8].

In order to involve the prospective users and interested parties in the early development of the project, we have initiated a study to identify their requirements, opinions and preferences. Appendix B describes the insights we have gathered from experts on the possibility of developing a clinical diagnosis decision support system for diagnosing rare diseases.

### 5.3 Sources of information for the rare disease project

In last few years, research on rare disorders has gained in importance. Researchers and clinicians share the data they collect by studying and treating rare diseases, patient organizations facilitate the dissemination of the collected information, and clinical trials for rare diseases are ongoing.

There are numerous rare disease patient organizations with the general common aim of improving the quality of life of people living with rare diseases and raise awareness of the diseases' impact on the lives of patients and their families. These organisations could provide useful information on rare diseases. One of these organizations is EURORDIS, the European Organisation for Rare Diseases, an alliance of patient organisations and individuals active in the field of rare diseases, representing more than 434 rare disease organizations in over 43 countries. In Denmark, there is an alliance of 36 national rare disease organisations called Rare Disorders Denmark<sup>1</sup>.

#### 5.3.1 Text-based information sources

Detailed information on rare diseases is available through databases maintained by organizations, agencies and other institutions. The prototype developed by M. Andersen and H. Jensen uses one of these <sup>2</sup>. It is a database of over 7,000 rare diseases, provided by the Office of Rare Diseases Research, National Institutes of Health, and contains description of rare diseases with annotated references and further resources (such as other databases, clinical trials and research, or testing laboratories and clinics).

The National Organization for Rare Disorders (NORD) provides over 1,200 rare disease reports written in non-technical language, and includes disease-related information such as: symptoms, causes, therapies, links to patient organizations and other resources, synonyms, disorder subdivision, affected population, and related disorders.<sup>3</sup>

Orphanet is a freely accessible database of rare diseases containing information on more than 5,000 diseases.<sup>4</sup> Data included in the database consists of a general disease summary, prevalence, inheritance, age of onset, ICD-10 code, synonyms, and links to further information such as classifications,

---

<sup>1</sup>Sjaeldne Diagnoser Danmark [www.sjaeldnediagnoser.dk](http://www.sjaeldnediagnoser.dk)

<sup>2</sup>[rarediseases.info.nih.gov](http://rarediseases.info.nih.gov)

<sup>3</sup>Abstracts are free, but full-text reports require a processing fee. [www.rarediseases.org](http://www.rarediseases.org)

<sup>4</sup>[www.orpha.net](http://www.orpha.net)

publications in PubMed, clinics, diagnostic tests, patient organizations, orphan drugs, research projects and clinical trials. The orpha.net portal also provides an interface in which it allows users to enter medical signs and search for a matching disease (around 2,000 diseases have been indexed for this system), maintains a journal (Orphanet Journal of Rare Diseases<sup>5</sup>) and publishes a series of reports that may be of interest to our project (e.g. a report on the prevalence of rare diseases).

Around 550 rare diseases (along with a general description, synonyms, diagnosis, treatment and other resources for each disease) are listed by the About.com Rare Disease page<sup>6</sup>. The Swedish National Board of Health and Welfare<sup>7</sup> provides a database including 265 diagnoses described in Swedish and 88 in English, including information such as: description, occurrence, cause, heredity, symptoms, diagnosis and testing, treatment or interventions, national and regional resources, courses and organizations, research, information material, along with corresponding literature.

Moreover, as a result of a collaboration between the Scandinavian countries, a rare disease database for nearly 580 diseases was produced (the project is called rarelink<sup>8</sup>). Although primarily described in one of the Scandinavian languages, reports of rare diseases also include some descriptions in English. The information provided by their database contains links to external websites providing information related to a disease (in English, Danish, Swedish, Finnish, Norwegian or Icelandic) and it also provides the terms used in the different languages to denote each disease.

We intend to further investigate these resources of information and search for others<sup>9</sup> as well, and to update our findings through the project's website<sup>10</sup>.

### 5.3.2 Non-text-based information sources

Because many rare diseases present a specific dysmorphic characteristic, systems have been developed that aid the clinician in identifying a disease solely based on the morphological features of the patient [37]. Such systems range from presenting a set of relevant photos or 3D models, up to identification based on image or 3D models analysis. [54, 55, 56]

---

<sup>5</sup>Latest articles on rare diseases, open-access to all articles, downloadable in PDF. Around four articles appear each month. [www.ojrd.com](http://www.ojrd.com)

<sup>6</sup>[rarediseases.about.com](http://rarediseases.about.com)

<sup>7</sup>[www.socialstyrelsen.se/rarediseases](http://www.socialstyrelsen.se/rarediseases)

<sup>8</sup>[www.rarelink.no](http://www.rarelink.no)

<sup>9</sup>Other sources of rare disease information we identified are: a genetic conditions database from Genetics Home Reference <http://ghr.nlm.nih.gov/BrowseConditions>, OMIM - Online Mendelian Inheritance in Man <http://www.ncbi.nlm.nih.gov/omim>, the list of rare diseases provided by HON - Health On the Net Foundation <http://www.hon.ch/HONselect/RareDiseases/index.html>, and Wikipedia's pages on the category of rare diseases [en.wikipedia.org/wiki/](http://en.wikipedia.org/wiki/)

<sup>10</sup>[sites.google.com/site/rarediseasescdss/about-rare-diseases](http://sites.google.com/site/rarediseasescdss/about-rare-diseases)

## 5.4 System design

### 5.4.1 Querying interface

As the Internet users today are accustomed to search using extremely short queries, we could assume that physicians may prefer the integration in the system of this rapid back-and-forth interaction style that comes with short queries [13]. Doctors often use web search engines which support this style, so this type of querying may be more natural to the physician when using the diagnostic system. This can be complemented by automatic filling of queries based on a couple of given keywords [13].

The clinician often knows more about a patient than the data he introduces in the computer [27]. Thus, if the system would be able to anticipate the user's needs and make further suggestions based on the keywords he already introduced, the clinician would just have to filter the information based on what he already knows but maybe did not include in the initial query.

The clinician can also be assisted by a system which recommends synonyms or related terms extracted from medical codifications such as the NCI Metathesaurus<sup>11</sup> or ICD-10<sup>12</sup>. Moreover, because some of these classifications are available in multiple languages, and because it is possible to correct spelling mistakes in query entries, the system might be appropriate also for clinicians for which English is a second language.

It would also be interesting to study whether a time dimension could be introduced in the list of symptoms [17]. This time dimension could be related to the appearance, sequence, or duration of signs and symptoms, and could prove to be important in matching these findings with similar cases.

### 5.4.2 Information retrieval and processing

Given that the amount of medical information about rare diseases is rapidly increasing, the system's database must be frequently updated. A web based system is the easiest distribution solution when the CDSS has to be frequently updated [35]. This would allow the user to always have the latest version, with the latest medical information. Nevertheless, this requires a mechanism for an easy update of the database and the system needs to be designed with an ability to be maintained and extended [32].

Text mining and natural language processing are two techniques that are widely employed in medical data analysis and can be and have been used in CDSSs [28]. Extensive medical knowledge has already been codified using the Unified Medical Language System (UMLS), the International Statistical Classification of Diseases and Related Health Problems 10th Revision

---

<sup>11</sup>NCI Metathesaurus maps more than 3,500,000 terms from 76 sources into 1,400,000 biomedical concepts. [ncimeta.nci.nih.gov](http://ncimeta.nci.nih.gov)

<sup>12</sup>[www.who.int/classifications/icd/en/](http://www.who.int/classifications/icd/en/)

(ICD-10) or Logical Observation Identifiers Names and Codes (LOINC). Moreover, there are numerous software products and libraries that utilize these classifications [13].

The use of PICO (Patient/Problem, Intervention, Comparison, and Outcome) framework was suggested as a good technique for structuring clinical queries. It was shown that using the PICO framework can result in clinicians formulating better queries that are easier to interpret by the system. Having a well-formulated query is a crucial step in finding answers to clinical questions [13].

### 5.4.3 Evidence strength

Because resources on the Internet vary greatly in terms of quality, relevance and popularity, it might be possible to use this information in our database model but also when presenting relevant resources for clinicians. There have been suggestions of classifying medical resources under three categories, based on the strength of the evidence [13]. Under this classification:

A-level evidence is based on resources that are consistent, of good quality, resulted from randomized control trials or meta-analyses.

B-level evidence is based on resources that are inconsistent, of limited quality, resulted from randomized control trials or meta-analyses.

C-level evidence is based on resources that are not rigorous.

A further classification could account for the publisher of that resource. In the case of MEDLINE citations, the study type is encoded into the *Publication Type field*[13].

It might also prove viable to employ a classification of resources based on the type of language they use, for example, classifying resources as using expert or non-expert medical terms [57], thus making a distinction between resources targeting patients or health care specialists.

## Chapter 6

# Conclusion

The findings of this study will be used in the development of the clinical diagnosis decision support system for diagnosing rare diseases. The evidence collected during this report shows that such a CDDS can add value to the existing medical care by providing evidence-based patient-specific advice at the point of decision making.

Our findings show that rare diseases can be suitable as the object of a specialized diagnostic CDSS. On one hand, as some diseases are so rare, the physicians may find it difficult to remember the details pertaining to every disease. Moreover, the patients affected by these diseases sometimes suffer long delays before a diagnostic is pronounced. This is especially problematic for cases where early diagnosis is essential for the efficacy of treatment [38].

During our research, we identified the existence of several medical codification standards and libraries for accessing and processing them. Clinicians have access to complex search engines that search through exhaustive medical article databases (such as PubMed), as well as several databases with information on rare diseases. The goal of the project will be to connect and combine all of these separate systems, in order to reduce the amount of time necessary to search for rare diseases, which was shown to be the main obstacle in clinicians using electronic medical resources in practice.

This study was an important first step towards designing the system. We intend to continue investigating user perspectives and plan to further integrate them in the design of the system.

## **Acknowledgements**

We are very thankful to our supervisor Ole Winther for letting us work on this project, his assistance to the project and his contribution in revising this paper. We would also like to thank chief physician Henrik L. Jørgensen for his input to the project in general, and our colleagues Henrik G. Jensen and Michael Andersen for their support in letting us further develop their work. We are also grateful to Dorica Dan, board member of EURORDIS, for her input to the project's proposal.

## **Authors' contribution**

The authors (RD and PP) equally contributed to writing the report. PP drafted the manuscript. Both authors discussed and commented on the work for this report.

## **Bibliography methodology**

A literature search was carried out in September 2010 by the authors, using Google Scholar and Science Direct, including or pertaining to keywords such as "clinical decision support system", "computer-assisted decision making", "computer-assisted diagnosis", "medical information retrieval", "representation of clinical data", "medline", "pubmed", and "medical (health) informatics". The authors independently performed this search. The authors reviewed, rated and tagged together the full texts of the retrieved articles. The reference sections of the retrieved articles were searched for any further relevant articles. The first search concluded in the retrieval of around 70 articles (duplicates eliminated), from which around 60% were considered relevant for the project. After the retrieval of the referenced articles that seemed to be of interest, the total number of articles grew to around 160. The final report uses around 35% of these.

# Appendix A. Questions to be answered during system evaluation

Described in and taken from [19]:

## Laboratory testing

### For a prospective user:

- (1) Structure (Is the system wanted?) Is there a perceived need? Can the advice be made available at the right place and time?
- (2) Process (Is it pleasant to use?) Is the interface satisfactory: desired options available? clear pointers to system state? effective screen-keyboard layout? is system dialogue acceptable? Is the system fast enough?
- (3) Outcome (Does it say sensible things?) Do its conclusions seem as sound as those of a respected authority? Are the explanations adequate?

### Questions that an expert authority might wish to ask:

- (1) Structure (Is the system of good quality?) Is the source of knowledge appropriate? Is the knowledge representation appropriate? Are the hardware and software adequate?
- (2) Process (Does it reason appropriately?) Is the logic consistent and rigorous? Is system control defined and clearly represented? Is the method of handling uncertainty rigorous? Is it robust to irrelevant variations in input data?
- (3) Outcome (Does it draw safe and potentially valuable conclusions?) Can it detect cases which are beyond its margins? Does it make seri-



ous mistakes within its domain? Compared to current practice, how accurate are its judgements?

## Field evaluation

- (1) Structure (how well does the system fit into its intended environment?)  
Do users find the overall system acceptable? Is adequate data available to the system? Is the system's advice accessible to the users? How well does the system integrate with other sources of expertise in the intended location?
- (2) Process (what effects does the system have on the processes of health care?)  
How often is the system used appropriately; what prevents this? Which parts of the system are not used or are abused, and why? Does the system have an effect on the quality or completeness of data? Does the system have an effect on the health care process: consultation time, length of stay number and types of investigations ordered, treatments used quality of users' decisions during and after use (education)? Does the system have a subjective effect on: users' job satisfaction, perceived responsibilities patients and administrators' perceptions of the interaction?
- (3) Outcome (does use of the system have an effect on outcome measures?)  
Does use of the system have an effect on individual patient morbidity? Does use of the system have an effect on population morbidity? Does use of the system have an effect on mortality? Are there any unexpected or significant side effects? What are the overall cost/benefit ratios associated with use of the system?

## Appendix B. Stakeholders' perspectives in the context of a CDSS diagnosing rare diseases

In a dialogue with Henrik L. Jørgensen, MD, PhD (chief physician at the Department of Clinical Biochemistry, Bispebjerg University Hospital, Denmark) and Dorica Dan (member of the Board of Directors at EURORDIS and president of the Romanian National Alliance for Rare Diseases), we found that the attitude towards such a system is positive and, conditioned by the improvement of the system, there is hope that it will prove useful to physicians as well as patients.

Following are their assessments on the idea of a project for developing a clinical decision support system for diagnosing rare diseases.

Dorica Dan: "It was a very pleasant surprise to find out about the existence of such a project to develop a support system for diagnosing rare diseases, given a list of symptoms of the patient. We always thought about it as well. There is one such option at the Orphanet portal, but the operation should be improved. We surely would want to be partners in such a project."

Henrik L. Jørgensen: "Overall, I think it is a very good idea with a system, which could suggest diagnoses of rare diseases to physicians. Many of these diseases are indeed very rare and the aggregated probability of meeting even one of them during a lifetime as physician is very low."

Mr. Jørgensen was involved from early on in the project researched by Henrik G. Jensen and Michael Andersen, and provided important input to their prototype. He now reflects on the outcome: "Michael and Henrik's system represents a promising prototype. I submitted 5 patient cases to them and in three of these, the correct diagnosis was among top 20. This hitrate obviously needs to be improved but it is a good start." He continues with a few recommendations for the future development of the system: "For the system to be of use in an everyday clinical setting, it must, of course, be

easy to use. The input should consist of symptoms reported by the patient, observations made by the physician as well as any additional information such as results of blood tests, X-rays etc. If the system can provide helpful suggestions as to the diagnosis, I am sure that it will be of interest both to general practitioners and to specialists alike.”

We intend to continue the investigation of users’ and stakeholders’ perspectives and plan to iteratively develop the system taking into account their needs.

However, the clinical setting in which such a system could be used in the future should be taken into consideration in the development process, and we have briefly studied two clinical settings in different countries in Europe (Denmark and Romania). Although we did not pursue a detailed study of these cases, it is obvious that there are differences between different parts of the world that may affect the interaction with the system. The high level of development and integration of computerized resources in Denmark is a major advantage over a country such as Romania.

# Bibliography

- [1] B. C. Delaney, “Potential for improving patient safety by computerized decision support systems.,” *Family practice*, vol. 25, pp. 137–8, June 2008.
- [2] J. Wyatt, “Use and sources of medical knowledge,” *The Lancet*, vol. 338, pp. 1368–1373, Nov. 1991.
- [3] B. C. Delaney, “Evidence-based diagnosis in general practice: needs both robust evidence and sophisticated electronic health record systems.,” *Family practice*, vol. 26, pp. 239–40, Aug. 2009.
- [4] P. Ramnarayan and J. Britto, “Paediatric clinical decision support systems,” *Archives of Disease in Childhood*, vol. 87, p. 361, Nov. 2002.
- [5] C. Sneiderman, D. Demner-Fushman, M. Fiszman, N. Ide, and T. C. Rindfleisch, “Knowledge-based methods to help clinicians find answers in MEDLINE,” *JAMIA*, pp. 772–780, 2007.
- [6] J. J. Cimino, J. Li, S. Bakken, and V. L. Patel, “Theoretical, empirical and practical approaches to resolving the unmet information needs of clinical information system users.,” *Proceedings / AMIA ... Annual Symposium. AMIA Symposium*, pp. 170–4, Jan. 2002.
- [7] J. W. Ely, J. A. Osheroﬀ, M. H. Ebell, M. L. Chambliss, D. C. Vinson, J. J. Stevermer, and E. A. Pifer, “Primary care care with evidence : qualitative study,” *Primary Care*, vol. 324, no. March, pp. 1–7, 2002.
- [8] R. A. Jenders, J. A. Osheroﬀ, D. F. Sittig, E. A. Pifer, and J. M. Teich, “Recommendations for clinical decision support deployment: synthesis of a roundtable of medical directors of information systems.,” *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, pp. 359–63, Jan. 2007.
- [9] D. W. Bates, G. Kuperman, S. Wang, T. Gandhi, A. Kittler, L. Volk, C. Spurr, R. Khorasani, M. Tanasijevic, and B. Middleton, “Ten commandments for effective clinical decision support: making the practice

- of evidence-based medicine a reality,” *Journal of the American Medical Informatics Association*, vol. 10, no. 6, pp. 523–530, 2003.
- [10] P. T. Johnson, J. K. Chen, J. Eng, M. A. Makary, and E. K. Fishman, “A comparison of world wide web resources for identifying medical information.,” *Academic radiology*, vol. 15, pp. 1165–72, Sept. 2008.
- [11] H. Yu and D. Kaufman, “A cognitive evaluation of four online search engines for answering definitional questions posed by physicians.,” *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, vol. 339, pp. 328–39, Jan. 2007.
- [12] R. Steinbrook, “Searching for the right search—reaching the medical literature.,” *The New England journal of medicine*, vol. 354, pp. 4–7, Jan. 2006.
- [13] D. Demner-Fushman and J. Lin, “Answering Clinical Questions with Knowledge-Based and Statistical Techniques,” *Computational Linguistics*, vol. 33, pp. 63–103, Mar. 2007.
- [14] G. De Leo, C. LeRouge, C. Ceriani, and F. Niederman, “Websites most frequently used by physician for gathering medical information.,” *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, p. 902, Jan. 2006.
- [15] T. Kortteisto, M. Kaila, I. Kunnamo, P. Nyberg, A.-M. Aalto, and P. Rissanen, “Self-reported use and clinical usefulness of second-generation decision supporta survey at the pilot sites for Evidence-Based Medicine elec-tronic Decision Support (,” *Finnish Journal of*, vol. 1, no. 3, pp. 161–169, 2009.
- [16] H. Tang and J. H. K. Ng, “Googling for a diagnosis—use of Google as a diagnostic aid: internet based study.,” *BMJ (Clinical research ed.)*, vol. 333, pp. 1143–5, Dec. 2006.
- [17] E. Berner, G. Webster, A. Shugerman, J. Jackson, J. Algina, A. Baker, E. Ball, C. Cobbs, V. Dennis, and E. Frenkel, “Performance of four computer-based diagnostic systems,” *New England Journal of Medicine*, vol. 330, no. 25, p. 1792, 1994.
- [18] I. Sim, P. Gorman, R. A. Greenes, R. Haynes, B. Kaplan, H. Lehmann, and P. C. Tang, “Clinical decision support systems for the practice of evidence-based medicine.,” *Journal of the American Medical Informatics Association : JAMIA*, vol. 8, no. 6, pp. 527–34, 2001.
- [19] J. Wyatt and D. Spiegelhalter, “Evaluating medical expert systems: what to test and how?,” *Medical informatics = Médecine et informatique*, vol. 15, no. 3, pp. 205–17, 1990.

- [20] D. F. Sittig, A. Wright, J. A. Osheroff, B. Middleton, J. M. Teich, J. S. Ash, E. Campbell, and D. W. Bates, “Grand challenges in clinical decision support.,” *Journal of biomedical informatics*, vol. 41, pp. 387–92, Apr. 2008.
- [21] F. Sullivan and J. Wyatt, *ABC of health informatics*. 2006.
- [22] C. Carroll, P. Marsden, P. Soden, E. Naylor, J. New, and T. Dornan, “Involving users in the design and usability evaluation of a clinical decision support system.,” *Computer methods and programs in biomedicine*, vol. 69, pp. 123–35, Aug. 2002.
- [23] J. Kassirer, “A report card on computer-assisted diagnosis,” *New England Journal of Medicine*, 1994.
- [24] G. Kong, D.-L. Xu, and J.-B. Yang, “Clinical Decision Support Systems: a Review of Knowledge Representation and Inference under Uncertainties,” *International Journal of Computational Intelligence Systems*, vol. 1, no. 2, p. 159, 2008.
- [25] M. L. Graber and A. Mathew, “Performance of a web-based clinical diagnosis support system for internists.,” *Journal of general internal medicine*, vol. 23 Suppl 1, pp. 37–40, Jan. 2008.
- [26] A. X. Garg, N. K. J. Adhikari, H. McDonald, M. P. Rosas-Arellano, P. J. Devereaux, J. Beyene, J. Sam, and R. Haynes, “Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review.,” *JAMA : the journal of the American Medical Association*, vol. 293, pp. 1223–38, Mar. 2005.
- [27] E. Berner and T. J. L. Lande, *Overview of Clinical Decision Support Systems (updated version of Chapter 36 in Ball MJ, Weaver C, Kiel J (eds). Healthcare Information Management Systems, Third Edition)*, vol. 6, ch. updated ve, pp. 463–477.
- [28] D. Demner-Fushman, W. W. Chapman, and C. J. McDonald, “What can natural language processing do for clinical decision support?,” *Journal of biomedical informatics*, vol. 42, pp. 760–72, Oct. 2009.
- [29] S. Zhu, M. Reddy, and J. Yen, “Hypothesis-Driven Story Building Framework: Enhancing Iterative Process Support in Clinical Diagnostic Decision Support Systems,” in *AMIA Annual Symposium Proceedings*, vol. 2009, p. 740, American Medical Informatics Association, Jan. 2009.
- [30] Y. Denekamp and M. Peleg, “TiMeDDx—a multi-phase anchor-based diagnostic decision-support model.,” *Journal of biomedical informatics*, vol. 43, pp. 111–24, Feb. 2010.

- [31] H. Yu, M. Lee, D. Kaufman, J. W. Ely, J. A. Osheroﬀ, G. Hripsak, and J. J. Cimino, “Development, implementation, and a cognitive evaluation of a deﬁnitional question answering system for physicians.,” *Journal of biomedical informatics*, vol. 40, pp. 236–51, June 2007.
- [32] M. Peleg and S. Tu, “Decision support, knowledge representation and management in medicine,” *Methods Inf Med*, pp. 72–80, Jan. 2006.
- [33] C. P. Friedman, “Enhancement of Clinicians’ Diagnostic Reasoning by Computer-Based Consultation: A Multisite Study of 2 Systems,” *JAMA: The Journal of the American Medical Association*, vol. 282, pp. 1851–1856, Nov. 1999.
- [34] M. Lee, J. J. Cimino, H. Zhu, C. Sable, V. Shanker, J. W. Ely, and H. Yu, “Beyond Information Retrieval Medical Question Answering,” *Symposium A Quarterly Journal In Modern Foreign Literatures*, pp. 8–10, 2006.
- [35] E. P. Hoffer, M. J. Feldman, R. J. Kim, K. T. Famiglietti, and G. Barnett, “DXplain: Patterns of Use of a Mature Expert System.,” *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, pp. 321–5, Jan. 2005.
- [36] F. Macartney, “Diagnostic logic,” *British Medical Journal*, vol. 295, no. 6609, pp. 1325–1331, 1987.
- [37] R. M. Winter, M. Baraitser, and J. M. Douglas, “A computerised data base for the diagnosis of rare dysmorphic syndromes.,” *Journal of medical genetics*, vol. 21, pp. 121–3, Apr. 1984.
- [38] M. G. Bouwman, Q. G. A. Teunissen, F. A. Wijburg, and G. E. Linthorst, “Doctor Google’ ending the diagnostic odyssey in lysosomal storage disorders: parents using internet search engines as an efficient diagnostic strategy in rare diseases.,” Aug. 2010.
- [39] J. Wyatt, “Computer-based knowledge systems,” *The Lancet*, vol. 338, pp. 1431–1436, 1991.
- [40] B. Chaudhry, “Computerized clinical decision support: will it transform healthcare?,” *Journal of general internal medicine*, vol. 23 Suppl 1, pp. 85–7, Jan. 2008.
- [41] C. Pearce, S. Trumble, M. Arnold, K. Dwan, and C. Phillips, “Computers in the new consultation: within the first minute.,” *Family practice*, vol. 25, pp. 202–8, June 2008.
- [42] H. Varonen, T. Kortteisto, and M. Kaila, “What may help or hinder the implementation of computerized decision support systems (CDSSs): a

- focus group study with physicians,” *Family practice*, vol. 25, pp. 162–7, June 2008.
- [43] I. I. Siempos, A. Spanos, E. A. Issaris, P. I. Rafailidis, and M. E. Falagas, “Non-physicians may reach correct diagnoses by using Google: a pilot study,” *Swiss medical weekly : official journal of the Swiss Society of Infectious Diseases, the Swiss Society of Internal Medicine, the Swiss Society of Pneumology*, vol. 138, pp. 741–5, Dec. 2008.
- [44] E. Murray, B. Lo, L. Pollack, K. Donelan, J. Catania, M. White, K. Zapert, and R. Turner, “The impact of health information on the internet on the physician-patient relationship: patient perceptions,” *Archives of internal medicine*, vol. 163, pp. 1727–34, July 2003.
- [45] A. G. Crocco, M. Villasis-Keever, and A. R. Jadad, “Analysis of Cases of Harm Associated With Use of Health Information on the Internet,” *JAMA: The Journal of the American Medical Association*, vol. 287, pp. 2869–2871, June 2002.
- [46] D. F. Sittig, M. A. Krall, R. H. Dykstra, A. Russell, and H. L. Chin, “A survey of factors affecting clinician acceptance of clinical decision support,” *BMC medical informatics and decision making*, vol. 6, p. 6, Jan. 2006.
- [47] N. Shebl, B. Franklin, and N. Barber, “Clinical decision support systems and antibiotic use,” *Pharmacy World & Science*, vol. 29, pp. 342–349, Aug. 2007.
- [48] K. Monika, L. Jamy, M. Christine, N. David, and S. Sharon, “Development of a prototype clinical decision support tool for osteoporosis disease management: a qualitative study of focus groups,” *BMC Medical Informatics and Decision Making*, vol. 10, p. 40, Jan. 2010.
- [49] D. Short, M. Frischer, and J. Bashford, “Barriers to the adoption of computerised decision support systems in general practice consultations: a qualitative study of GPs’ perspectives,” *International journal of medical informatics*, vol. 73, pp. 357–62, May 2004.
- [50] J. Henderson, “Google Scholar: A source for clinicians,” *Canadian Medical Association Journal*, vol. 172, no. 12, pp. 1549–1550, 2005.
- [51] H. G. Jensen and M. Andersen, “Support decision system for diagnosing rare diseases using vector space model and medical text mining,” no. January, 2010.
- [52] G. Purcell, “What makes a good clinical decision support system,” *BMJ*, vol. 330, pp. 740–1, Apr. 2005.



- [53] K. Kawamoto, C. A. Houlihan, E. A. Balas, and D. F. Lobach, “Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success.,” *BMJ (Clinical research ed.)*, vol. 330, p. 765, Apr. 2005.
- [54] P. Hammond, T. J. Hutton, J. E. Allanson, L. E. Campbell, R. C. M. Hennekam, S. Holden, M. a. Patton, A. Shaw, I. K. Temple, M. Trotter, K. C. Murphy, and R. M. Winter, “3D analysis of facial morphology.,” *American journal of medical genetics. Part A*, vol. 126A, pp. 339–48, May 2004.
- [55] S. Boehringer, T. Vollmar, C. Tasse, R. P. Wurtz, G. Gillessen-Kaesbach, B. Horsthemke, and D. Wieczorek, “Syndrome identification based on 2D analysis software.,” *European journal of human genetics : EJHG*, vol. 14, pp. 1082–9, Oct. 2006.
- [56] P. Hammond, “The use of 3D face shape modelling in dysmorphology.,” *Archives of disease in childhood*, vol. 92, pp. 1120–6, Dec. 2007.
- [57] N. Grabar and S. Krivine, “Application of Cross-Language Criteria for the Automatic Distinction of Expert and Non Expert Online Health Documents,” *Artificial Intelligence in Medicine*, pp. 252–256, 2007.