Problems in Establishing the Medical Expert Systems CADIAG-1 and CADIAG-2 in Rheumatology

Gernot Kolarz and Klaus-Peter Adlassnig

CADIAG-1 and CADIAG-2 are medical expert systems with applications in rheumatology, gastroenterology, and hepatology. CADIAG-1 is based on a symbolic logic representation of medical relationships between symptoms, signs, or findings and diseases. Definite relationships (obligatory occurrence, confirming, and excluding) as well as uncertain relationships (frequent occurrence and not confirming) are applied to confirm or exclude diagnoses and to establish diagnostic hypotheses. CADIAG-2 employs fuzzy set theory and fuzzy logic to formalize medical entities and relationships. The medical concept of confirming or excluding diagnoses is identical to that of CADIAG-1, but diagnostic hypotheses are generated differently. Here, a documentation of medical relationships allowing gradual transitions from 'always' to 'never' for the frequencies of occurrence of symptoms with and from 'strong' to 'weak' for their strengths of confirmation for diseases leads to strongly or weakly supported diagnostic hypotheses in the actual case. Tests with 322 real patient cases from a rheumatological hospital, each including between 500 and 700 symptoms, signs, and findings, were carried out. The percentage of cases diagnosed correctly is about 80%. Problems and pitfalls that became apparent in the evaluation of the cases are shown and discussed.

INTRODUCTION

For more than two decades various methods for computer-assisted diagnosis in medicine have been developed. They are often based on statistical approaches like Bayesian models,\(^1\)\(^-\)\(^3\) discriminant analysis,\(^4\)\(^-\)\(^5\) factor analysis,\(^6\) and cluster analysis.\(^7\) More recently, artificial intelligence principles have been applied in several computer-assisted medical expert systems.\(^8\)\(^-\)\(^1\)\(^1\) These approaches seem to be more successful, especially in extended medical areas in which up to several hundred diagnoses and several thousand symptoms, signs, and findings are considered.

From the Ludwig Boltzmann Institute for Rheumatology and Focal Diseases, Kaiser-Franz-Ring 8, A-2500 Baden, Austria, the Computer Science Division, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California 94720, and the Department of Medical Computer Sciences, University of Vienna, Garsnitzgasse 13, A-1090 Vienna, Austria.
Since 1968, two computer-assisted diagnostic systems—CADIAG-1 and CADIAG-2—have been developed at the Department of Medical Computer Sciences, University of Vienna. These systems include some small as well as large branches of internal medicine, such as rheumatology, hepatology, pancreatic diseases, gall bladder and bile duct diseases, and colon diseases.

The aim of this paper is to show the problems and pitfalls that occur using these systems in rheumatology where several medical specialties overlap, such as internal medicine, orthopedics, physical and psychosomatic medicine, urology, and others. Furthermore, it is usually elderly patients who suffer from rheumatic diseases; therefore, multimorbidity in these cases often complicates the diagnostic process.

METHODS

CADIAG-1/RHEUMA

In CADIAG-1, 187 diagnoses and 1,213 symptoms, signs, and findings have been included in the rheumatological differential diagnostic group CADIAG-1/RHEUMA, and relationships between them have been documented. The symptoms, signs, and findings contain the case history, the physical status, the X-ray and lab results, and the histological findings. Detailed descriptions of the computer documentation of patient data have been published.

The medical knowledge that is the basis of the CADIAG-1 system is formed by relationships in the following categories:

1. relationships between symptoms, signs, or findings and diseases:
   - **OC**: a symptom, sign, or finding shows *obligatory occurrence* with a disease, and if it occurs it is *confirming* for that disease.
     Example 1
     The X-ray finding “endoprothesis of the knee” is *obligatory occurring* and *confirming* for the diagnoses “arthroplasty of the knee.”
   - **FC**: a symptom, sign, or finding shows *facultative occurrence* with a disease, but if it occurs it is *confirming* for that disease.
     Example 2
     The lab result “intracellular uric acid crystals in joint effusion” is *facultative occurring* yet *confirming* for the diagnosis “gout.”
   - **ON**: a symptom, sign, or finding has *obligatory occurrence* with a disease but does not confirm it.
     Example 3
     The clinical finding “HEBERDEN’s nodes” is *obligatory occurring* and *not confirming* for the diagnosis “HEBERDEN’s arthrosis.”
   - **FN**: a symptom, sign, or finding is both *facultative occurring* with a disease and *not confirming* for that disease.
Example 4
The lab finding “elevated ESR” is facultative occurring and not confirming for the diagnosis “rheumatoid arthritis.”

\( E: \) a symptom, sign, or finding excludes a disease.

Example 5
The lab finding “WAALER ROSE titre 1:128” excludes the diagnosis “sero-negative rheumatoid arthritis.”

\( \sim \): the symptom–disease relationship is unknown or unspecific.

2. relationships among symptoms, signs, or findings, among diseases, and between symptom combinations and diseases:
only OC, FC, ON, and E relationships are allowed.

After documenting all known relationships for the considered rheumatological diseases, a special computer program is applied for some preliminary calculations. For every symptom, sign, or finding, the frequency of \( FN \) relationships to diseases is calculated to obtain a degree of ambiguity of that symptom, sign, or finding. Symptoms, signs, or findings with a low degree of ambiguity are used to compute unique symptom patterns in the medical documentation.

The diagnostic process is then performed in the following way: Given a certain symptom pattern, an attempt is made to infer confirmed diagnoses, excluded diagnoses, and diagnostic hypotheses. Confirmed diagnoses can be determined if present symptoms, signs, or findings show a relationship of the OC or FC category to a particular disease. Excluded diagnoses can be established from present symptoms, signs, or findings and E relationships, as well as from definitely absent symptoms, signs, or findings with OC or ON relationships. Diagnostic hypotheses are generated if unique symptom patterns match the patient’s present symptoms, signs, or findings. The computer system is able to display the line of reasoning for its inferred diagnoses, thus making the diagnostic process comprehensible. The generation of proposals for further useful investigations and the subsequent input of the results of such investigations enable the physician to perform the diagnostic process iteratively. Because of that, he can confirm or eliminate diagnostic hypotheses step by step.

**CADIAG-2/RHEUMA**

When testing the CADIAG-1 system, it became soon apparent that it would be advantageous to have gradual transitions between adjacent medical concepts. This concerns transitions from normal to pathological, from facultative to obligatory, and from not confirming to confirming. Therefore, in a second step, CADIAG-2 was developed (it contains the same symptoms, signs, or findings, and diseases as CADIAG-1), using fuzzy set theory as a mathematical basis that provides the possibility of defining inexact medical entities as fuzzy sets. In order to determine the relationships

\* At some places in this paper, the term symptom is considered to be synonymous with the terms sign and finding.

between symptoms, signs, or findings and diseases, the frequency of occurrence of a symptom, sign, or finding with a disease and the strength of confirmation of a symptom, sign, or finding for a disease are taken into consideration. Examples for the representation of medical knowledge in CADIAG-2/RHEUMA are shown below:

Example 6
IF (the patient has increased uric acid in serum)
THEN (the diagnosis may be gout)
WITH (the frequency of occurrence value of hyperuricaemia in gout is often [.75], and the strength of confirmation value of hyperuricaemia for gout is weak [.20]).

Example 7
IF (the patient shows tophi)
THEN (the diagnosis is gout)
WITH (the frequency of occurrence value of tophi with gout is seldom [.25], and the strength of confirmation value of tophi for gout is always [1.00]).

Two ways are offered to document the relationships between symptoms, signs, or findings and diseases: (1) linguistic documentation of the relationships by expert physicians who apply terms such as always, almost always, very often, very strong, . . . medium, . . . and never to describe the respective relationships between the medical entities;12,24 (2) automatic calculation of the relationships by evaluating a patient data base with already diagnosed patients and determination of both the frequencies of occurrence (relative frequency of occurrence of symptom S with disease D—that is, F(S/D)) —and the strengths of confirmation (relative frequency of occurrence of disease D with symptom S—that is F(D/S))—attributed to symptoms, signs, or findings and diseases.12,14,24

The linguistic documentation of the relationships makes it possible to document medical knowledge directly from medical textbooks. The automatic calculation of relationships can be applied to offer a basis for the documentation of diseases or to check already documented medical knowledge.

The representation and documentation of the other relationships considered in CADIAG-2 (symptom—symptom, disease—disease, and symptom combination—disease relationships) is conducted in an analogous way. An example for a symptom combination—disease relationship is the following:

Example 8
IF (the patient shows low back pain, and
a limitation of motion of the lumbar spine, and
a diminished chest expansion, and
the patient is male, and
is between 20 and 40 years of age)
THEN (the diagnosis may be ankylosing spondylitis)
WITH (the frequency of occurrence value of the above combination with ankylosing spondylitis is very often [.90] and the strength of confirmation value of the combination for ankylosing spondylitis is very strong [.80]).
The diagnostic process in CADIAG-2 is carried out as follows: Given a certain symptom pattern, fuzzy grades of membership of the patient to diagnoses are calculated by means of composing fuzzy relationships (patient–symptom relationships, patient–symptom combination relationships, symptom–symptom relationships, symptom–disease relationships, symptom combination–disease relationships, and disease–disease relationships). Similar to CADIAG-1, the results are confirmed and excluded diagnoses as well as diagnostic hypotheses. Explanations for the established diagnoses and proposals for further investigations are also displayed.

RESULTS

Documentation

The differential diagnostic groups of rheumatic diseases contain 187 diseases, among them joint diseases, diseases of the spinal column, diseases of soft tissue and connective tissue system, diseases of cartilage and bone, systemic diseases with facultative manifestations in the locomotor apparatus, and regional pain syndromes.

The categorization of the 1,213 symptoms, signs, and findings applied in order to document the 187 diseases is as follows: 278 symptoms of patient’s history, 488 signs from physical examination, 270 laboratory test results (from about 70 lab tests), 15 findings of causative agents, 53 biopsy and histology findings, 89 X-ray findings, 1 ECG finding, and 19 concomitant diseases.

Medical Knowledge Documentation of CADIAG-1/RHEUMA. The documentation of symptoms, signs, or findings and their relationships to the considered diseases is completed. Altogether, there are 15,814 (100%) relationships documented between symptoms, signs, or findings and diseases, among them 506 (3.2%) definite relationships (i.e., of the categories OC, FC, ON, and E) and 15,308 (96.8%) uncertain relationships (i.e., of the category FN).

Because of the high proportion of uncertain relationships, it is easy to understand that confirmed diagnoses are very infrequent. To overcome this problem for several diseases, definite symptom combinations have been established using worldwide accepted diagnostic criteria, such as the ARA criteria for rheumatoid arthritis28 or the Jones criteria for rheumatic fever.29 Altogether, 32 symptom combinations, including 13 that are obligatory combinations for diagnosing diseases, are contained in CADIAG-1/RHEUMA. Furthermore, 1,010 relationships among symptoms, signs, and findings themselves that capture hierarchical dependencies between these patient data as well as definite exclusions between them were also documented. Finally, 1,189 relationships among diseases (taxonomy, exclusions) were documented as well.

Medical Knowledge Documentation of CADIAG-2/RHEUMA. The documentation for CADIAG-2/RHEUMA is also completed. Frequency of occurrence relationships were documented 15,763 (100%) times and relationships of the category strength of confirmation 15,745 (100%) times. Always (obligatory) occurrence was selected 42 times and never (exclusion) occurrence 333 times. For the strength of confirmation, always (confirming) occurred 70 times and never (exclusion; simultaneously to never occurrence) 333 times. The remaining part are uncertain relationships; the figures are 15,388 (97.6%)
relationships for the frequency of occurrence and 15,342 (97.4%) relationships for the strength of confirmation.

Additionally to the symptom combinations defined in CADIAG-1/RHEUMA, several combinations that strongly indicate, but not confirm, some diseases were included in CADIAG-2/RHEUMA (9 additional combinations).

Relationships among symptoms, signs, or findings and among diseases are documented identically to those in CADIAG-1/RHEUMA.

**Patient's Data.** Since 1981 more than 3,500 patients of a rheumatologic unit have been documented with their diagnoses, case histories, physical examination results, and lab and X-ray findings. The number of symptoms, signs, or findings collected for each patient varies from 500 to 700. The documented rheumatological diagnoses vary from 1 to 8 per patient. The number of the diagnoses therefore exceeds the number of the documented patients. In Table 1 it can be seen that the greatest number of patients suffered from "degenerative joint disease," from "soft tissue disease," or from "disorders of the vertebral column."

Patients with one of the following diagnoses were tested in the two computer-assisted diagnostic system CADIAG-1/RHEUMA and CADIAG-2/RHEUMA: "Rheumatoid arthritis, SJÖGREN's syndrome, ankylosing spondylitis, systemic lupus erythematosus, gout," and "progressive systemic sclerosis." These diagnoses were selected for testing because well-defined confirming symptom combinations are available in both systems to establish these diseases as diagnoses. The percentage of patients diagnosed correctly, in comparison to the clinical diagnoses, is shown in Table 2.

All cases in which the computer diagnoses did not match the clinical diagnoses, or where a computer diagnosis could not be established, were studied for the reason of the failure.

**Table 1. Number of Documented Patients According to the Most Frequent Diagnoses in a Rheumatological Hospital**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N documented patients</th>
<th>N documented case histories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osteoarthritis</td>
<td>2,674</td>
<td>2,741</td>
</tr>
<tr>
<td>Arthroplasty</td>
<td>351</td>
<td>376</td>
</tr>
<tr>
<td>Degenerative disease of the vertebral column</td>
<td>2,252</td>
<td>2,343</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>193</td>
<td>212</td>
</tr>
<tr>
<td>Fibrositis</td>
<td>2,214</td>
<td>2,240</td>
</tr>
<tr>
<td>Gout</td>
<td>46</td>
<td>47</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>191</td>
<td>223</td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>Psoriatic arthropathy</td>
<td>21</td>
<td>28</td>
</tr>
<tr>
<td>Others</td>
<td>502</td>
<td>545</td>
</tr>
</tbody>
</table>
Table 2. Percentage of Correctly Diagnosed Cases by the Medical Expert Systems

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N cases</th>
<th>Absolute</th>
<th>In percentage</th>
<th>Reason for failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rheumatoid arthritis</td>
<td>223</td>
<td>200</td>
<td>89.7%</td>
<td>Treatment</td>
</tr>
<tr>
<td>Sjögren’s syndrome</td>
<td>10</td>
<td>9</td>
<td>90.0%</td>
<td>Treatment</td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>32</td>
<td>29</td>
<td>90.6%</td>
<td>Case history, no typical X-ray finding</td>
</tr>
<tr>
<td>Systemic lupus erythematosus</td>
<td>6</td>
<td>4</td>
<td>66.7%</td>
<td>Treatment</td>
</tr>
<tr>
<td>Gout</td>
<td>47</td>
<td>18</td>
<td>38.3%</td>
<td>Treatment, case history</td>
</tr>
<tr>
<td>Progressive systemic sclerosis</td>
<td>4</td>
<td>3</td>
<td>75.0%</td>
<td>Overlapping syndrome</td>
</tr>
</tbody>
</table>

* Diagnoses that could be established by confirming symptom combinations implemented in both systems were tested.

Reasons for Failure of the Medical Expert Systems

The *case history* of a patient is often incomplete and sometimes even wrong. This depends partly on the development of medicine, since diseases that occurred in a patient, e.g., 20 years ago may not have been diagnosed accurately but would be important if they were true. For example, many patients tell about rheumatic fever in their case history, but they did not get antibiotics or cortisone. If cardiac symptoms are absent, one might assume that this kind of polyarthritis could have been some sort of reactive arthritis, which might be without any importance for the acute disease. Another example is the circumstances at the onset of the disease. Human beings are inclined to reason every event. Patients therefore very often connect a special event with the onset of their disease. On the other hand, causally connected symptoms may be forgotten. For traumatic spondylopathy diagnosed by X-ray examination, where the old fracture could be diagnosed, only 83% of the patients mentioned a former trauma. Even in patients with hip arthroplasty, only 75% indicated the former surgical treatment on the questionnaire of patient’s history. Therefore, symptoms of the case history could no longer be linked to diagnoses by ON relationships in CADIAG-1 or always-occurrence relationships in CADIAG-2.

Another difficulty is the *definition of normal borderlines* of physical findings. Even if normal ranges for a certain sign were accepted worldwide, problems might arise with signs that change with the age of the patient. An example is the mobility of the lumbar spine. Patients over 60 years often show restricted mobility in comparison to healthy younger people. Therefore, normal ranges vary not only for both sexes but also for different ages, body weights, and other variables. Thus, physical status as well as lab findings have to be indicated as either normal or pathological by the physicians, or
normal ranges evaluated by the computer have to consider different physiological variables. Here, the CADIAG-2 approach by applying fuzzy sets that allow a gradual transition from normal to pathological ranges has some advantages.

The documentation of X-ray findings is as complex as the result of a consultation of other medical specialists. If the computerized diagnostic system contained the complete body of medical knowledge, one would have to document only signs, e.g., "narrowing of the joint space, subchondral sclerosis, osteophytes," but not the X-ray diagnosis "osteoarthritis," or "redness of the throat, swollen tonsils with purulent depositions," but not the laryngologic diagnosis of a "bacterial tonsillitis." Because, in the various rheumatic diseases, nearly every organ can show signs (e.g., central nervous signs, pulmonary signs, or renal signs in connective tissue diseases; severe intestinal signs in "arthritis with CROHN's disease"; hematological signs in "arthritis with hemophilia"; etc.), such a diagnostic system would have to contain almost the entire medical knowledge. If all possible symptoms, signs, and findings in all possible rheumatic diagnoses were documented, such a system would probably be too large to handle. Therefore, the only possibility to overcome this problem is to document the diagnoses of other specialists as symptoms for the rheumatologic system. This method implies that some potentially important findings may not be documented. For the X-ray findings a compromise was chosen: As X-ray findings of the skeleton are documented with their symptoms and signs, X-ray findings of other organs are documented with their diagnoses.

The classification of diagnoses is not satisfying either. Various systems classifying diagnoses were established, where the number of rheumatic diagnoses varies from about 200 to 1,330. Furthermore, very often there is an overlap of diagnoses. The best example is the overlapping syndromes in connective tissue diseases. Patients with overlapping syndromes often could not be diagnosed by the computer-assisted diagnostic systems. The symptom "scleroderma," for example, excludes the diagnosis of "systemic lupus erythematosus." A patient in whom, on the one hand, "haemolytic anaemia, positive LE cells" and "proteinuria" and, on the other, "sclerodactily, impaired oesophageal motion" and "lung fibrosis" are found could not be diagnosed by the computer for there are symptoms present of both "systemic lupus erythematosus" and "progressive systemic sclerosis."

The greatest problem of all is the influence of the therapy. Drug therapy, especially, can modify the course of a disease in several ways. On the one hand, symptoms of inflammation may for example be diminished or abolished by the use of corticosteroids, so that in "rheumatoid arthritis" one does not necessarily need to find "elevated ESR, swollen and tender joints," and "subfebrile temperature." On the other hand, adverse effects like "peptic ulcer" or "gastrointestinal bleeding" may occur and thus confuse the clinical and lab symptomatology. In similar ways this is true for surgical and physical treatment. Despite the various treatments, usually enough symptoms remain to correctly diagnose a disease in a patient. In some cases, however, the computer-assisted systems have not been successful.

Patients with rheumatic diseases, especially with degenerative joint diseases, are often of middle or older age. In these patients various concomitant diseases are common that are not connected to the rheumatic disease. The symptoms of, e.g., "hypertension, arterial sclerosis, diabetes" may confuse the clinical picture and the lab findings of the rheumatic disease that should be diagnosed. This problem is partly solved by using the
CADIAG-2 system where the medical strength of confirmation of a symptom or sign for the disease is indicated in detail.

**DISCUSSION**

The majority of existing medical diagnostic systems in rheumatology is confined to few diagnoses, which are more or less well defined. Our system includes the whole spectrum of rheumatic diseases with more than 18,000 relationships (CADIAG-1/RHEUMA) and 33,000 relationships (CADIAG-2/RHEUMA), respectively, between symptoms, signs, or findings and diseases; among symptoms, signs, and findings; among diseases; and between symptom combinations and diseases. But only 270 of all relationships could be found in the category confirming (strength of confirmation = always); a further 300 were in the category obligatory occurrence (frequency of occurrence = always); and about 2,198 are excluding relationships (frequency of occurrence = never and strength of confirmation = never). More than 15,000 (CADIAG-1/RHEUMA), i.e., about 83%, and 30,000 (CADIAG-2/RHEUMA), i.e., about 90%, relationships documented are uncertain relationships. Especially for rheumatic diseases, which are not defined by generally accepted diagnostic criteria, it was therefore desirable to differentiate between symptoms, signs, and findings with respect to their sensitivity (frequency of occurrence) and medical specificity (strength of confirmation) to a certain diagnosis. This led to the development of CADIAG-2, where a so-called typical symptom, sign, and finding for a disease is taken into consideration more than another, although both have only $FN$ relationships for the disease under consideration in CADIAG-1.

The overall percentage of 81.7% correct diagnoses in the 322 tested cases seems to be satisfying if we take into consideration that, especially with "gout," the specific treatment given for several years may abolish all typical symptoms.

The problems occurring are due to practical work, where real patients show a lot of interfering diseases. The development of CADIAG-2 is capable of overcoming these problems to some extent. But one has to be aware that a computer-assisted diagnostic system will not be as successful as an experienced rheumatologic specialist. But it can be helpful for teaching systems where students can imitate diagnostic procedures, and it may be of value for general practitioners in patients with rare rheumatic diseases.

**ACKNOWLEDGMENTS**

This work was supported by a Max Kade Fellowship for Scientific Research and was also supported in part by National Science Foundation NSF-IST Grant 8320416 and NASA Grant NCC2-275.

**REFERENCES**


