KNOWLEDGE ACQUISITION STUDY AND ACCURACY RATE EVALUATION FOR CADIAG-2/RHEUMA WITH 308 CLINICAL CASES

Harald Leitich[1], Klaus-Peter Adlassnig[1], Gernot Kolarz[2]

[1]Department of Medical Computer Sciences
(Director: Univ.Prof. Dr. G. Grabner)
University of Vienna, Garnisongasse 13, A - 1090 Vienna, Austria
and
[2]Ludwig Boltzmann Institute for Rheumatology and Focal Diseases
(Director: Univ.Prof. Dr. N. Thumb)
Kaiser-Franz-Ring 8, A - 2500 Baden, Austria

Abstract

CADIAG-2/RHEUMA is a medical expert system assisting in the differential diagnosis of rheumatic diseases. The aim of this study was to establish a set of criteria for diagnosing definite rheumatoid arthritis (RA) that provides optimal accuracy and to implement this set of criteria as an IF-THEN rule for application in CADIAG-2/RHEUMA. First, two different sets of criteria for the classification of definite RA described in medical literature were implemented and their respective diagnostic accuracies were evaluated with 154 patients suffering from RA and 154 control subjects. Second, that set of criteria which had performed best became the starting point for establishing an improved set of diagnostic criteria that eventually reached an accuracy of 88.7% (81.8% sensitivity and 95.5% specificity). This improvement was possible by combining literature definition of RA with specific clinical experience of a rheumatology expert.

1. Introduction

CADIAG-2/RHEUMA is a medical expert system designed to assist in establishing differential diagnoses of rheumatic diseases. Its theoretical background and integration into a medical information system were described in [1,2,3,5]. The expert system's accuracy has been evaluated before, reaching 93.7% sensitivity by evaluating 426 patients with rheumatoid arthritis, gout, ankylosing spondylitis, psoriatic arthritis, Sjögren's disease, systemic lupus erythematosus, Reiter's disease, and systemic sclerosis [2].

This study is focused on the diagnosis of rheumatoid arthritis (RA). Three different sets of RA diagnostic criteria were used and applied to RA patients and a control group with non-RA rheumatic diseases. All three sets were based on RA classification criteria edited by The American Rheumatism Association (ARA, now The American College of Rheumatology) [4,5]. In CADIAG-2/RHEUMA, each set of diagnostic criteria was implemented as one IF-THEN rule. The three established rules are:

- rule 1 was implemented according to the revised 1958 ARA criteria for the classification of definite RA [6];
- rule 2 was implemented according to the revised 1987 ARA criteria for the classification of (definite) RA [4];

- rule 3 combined both literature definition and specific clinical experience of a rheumatology expert: As rule 1, it was implemented according to the revised 1958 ARA criteria for the classification of definite RA; however, several criteria were changed and redefined by the rheumatology expert.

The aims of the study were: (a) which of the two initial sets of RA diagnostic criteria performed best; (b) whether the best performing set of criteria could still be improved; and (c) whether splitting both RA patients and control subjects into subsets (according to disease stages, disease characteristics, and concomitant diseases) would give a more detailed picture of the expert system's performance.

2. Patient data

All 154 RA patients and 154 control subjects of this study underwent treatment in a 140-bed hospital for rheumatic diseases in Baden/Austria. Only adults with disease onset after age 16 were included. The mean ages of RA and non-RA patients were similar, whereas sex percentages differed because of a larger number of male control patients.

RA patients: 150 of the 154 RA patients had a confirmed clinical diagnosis of RA. The remaining 4 patients were diagnosed as suspected RA cases. (At time of data collection, a more specific diagnosis was not possible; they were, however, confirmed as definite RA cases later.) The confirmed RA patients were additionally subdivided into the following groups:

- disease stage (The disease staging used in this study is based on radiographic findings and was introduced by Steinbrocker et al. [7]);
- presence or absence of rheumatoid factor (seropositivity/seronegativity); and
- concomitant rheumatic diseases with no relation to RA according to the following categorization: (1) concomitant diseases of the vertebral column; and (2) concomitant osteoporosis.

Control subjects: Any patient with a rheumatic disease other than RA was designated a control subject. The clinical diagnoses represent a cross-section of patients that were treated in the hospital mentioned above.

3. Method

The results shown in Tables 1–5 were obtained by comparing CADIAG-2/RHEUMA's diagnostic results with the available confirmed clinical diagnoses. In addition, it has to be mentioned that a CADIAG-2/RHEUMA diagnosis was taken as established if it was either a confirmed diagnosis or a diagnostic hypothesis with a degree of confirmation of at least 0.5 (cf., [1,3]).

4. Results

4.1. Results obtained with rules 1 and 2 (based on literature definition)

True positive results (sensitivity) with RA patients: As shown in Table 1, rule 1 performed best reaching a sensitivity of 75.3%. Tables 2–4 show the different diagnostic results obtained in the RA subgroups. Cases with early disease stage (stage 1), cases with seronegative RA, and cases with concomitant diseases of the vertebral column tended to cause a lower sensitivity.

False positive results (100%−specificity rate) with control subjects: As shown in Table 5, rule 1 performed best reaching a specificity of 87.7%. A substantial number of patients with psoriatic arthritis and systemic lupus erythematosus yield incorrect results, a fact which lead to further developments, as is described in Section 4.2.

Table 1: True positive results (sensitivity) obtained by the literature-based rules 1 and 2 in patients with rheumatoid arthritis (RA).

diagnosis	total number of patients	rule 1 (ARA 1958 definite)	rule 2 (ARA 1987 definite)
suspected RA	4	1	1
seropositive RA, stage 1	9	7	6
seronegative RA, stage 1	19	10	10
seropositive RA, stage 2	25	22	21
seronegative RA, stage 2	26	17	16
seropositive RA, stage 3	28	23	25
seronegative RA, stage 3	20	18	15
seropositive RA, stage 4	17	13	12
seronegative RA, stage 4	6	5	5
total number of diagnoses	154	116	111
sensitivity rates		75.3%	72.1%

Table 2: True positive results (sensitivity) obtained by the literature-based rules 1 and 2 in RA patients, disease stages 1–4.

diagnosis	total number of patients	rule 1 (ARA 1958 definite)	rule 2 (ARA 1987 definite)
RA, stage 1	28	17 (61%)	16 (57%)
RA, stage 2	51	39 (76%)	37 (72%)
RA, stage 3	48	41 (85%)	40 (83%)
RA, stage 4	23	18 (78%)	17 (74%)

Table 3: True positive results (sensitivity) obtained by the literature-based rules 1 and 2 in seropositive and seronegative RA patients.

diagnosis	total number of patients	rule 1 (ARA 1958 definite)	rule 2 (ARA 1987 definite)
seropositive RA	79	65 (82%)	64 (81%)
seronegative RA	71	50 (70%)	46 (64%)

Table 4: True positive results (sensitivity) obtained by the literature-based rules 1 and 2 in RA patients with and without concomitant spinal diseases.

diagnosis	total number of patients	rule 1 (ARA 1958 definite)	rule 2 (ARA 1987 definite)
no concomitant diseases	74	59 (80%)	54 (73%)
diseases of the vertebral column	50	36 (72%)	34 (68%)
osteoporosis	26	20 (77%)	22 (85%)

Table 5: False positive results (100%−specificity rate) obtained by rules 1 and 2 in control subjects.

diagnosis	total number of patients	rule 1 (ARA 1958 definite)	rule 2 (ARA 1987 definite)
osteoarthrosis	44	2	0
gouty arthritis	32	0	4
ankylosing spondylitis	30	2	0
psoriatic arthritis	20	10	10
joint tuberculosis	4	1	1
other joint infections	4	1	1
Reiter's disease	4	0	0
systemic lupus erythematosus	4	2	3
systemic sclerosis	4	0	0
polymyositis	3	0	1
chondrocalcinosis	3	1	0
polymyalgia rheumatica	2	0	0
total number of diagnoses	154	19	20
specificity rates		87.7%	87.0%

4.2. Development of an improved rule 3 (based on literature definition and clinical experience)

To improve CADIAG-2/RHEUMA's performance, clinical experience of a rheumatology expert was needed to modify the established rules. Rule 1 was selected for further improvement because of its obtained high sensitivity and specificity. Its criteria were consecutively changed to reach higher rates of sensitivity and specificity. Finally, the problem was successfully approached in two different ways:

(a) redefinition of some diagnostic criteria

The symptom "morning stiffness" was redefined in a more restrictive way, but had to last for only 30 minutes instead of 60 minutes.

The sign "symmetrical joint swelling", which had to be observed by a physician, was remodelled to "symmetrical joint involvement", with the additional inclusion of patient history data.

(b) addition of some further exclusion criteria

To avoid false positive results in cases of psoriatic arthritis, an exclusion in case of present psoriasis was added to rule 1. This exclusion prevents the diagnosis of definite RA if there is sufficient evidence that a patient might actually suffer from psoriatic arthritis.

4.3. Diagnostic results obtained with the improved rule 3 (based on literature definition and clinical experience)

All improvements lead to a definite rule for RA which showed a sensitivity of 81.8% and a specificity of 95.5%, thus reaching a total accuracy of 88.7% (means of sensitivity and specificity rates) as is shown in Table 6.

Table 6: Sensitivity, specificity, and accuracy rates obtained by rules 1 and 2 and improved rule 3

	rule 1 (ARA 1958 definite)	rule 2 (ARA 1987 definite)	rule 3 (improved rule 1)
sensitivity	75.3%	72.1%	81.8%
specificity	87.7%	87.0%	95.5%
accuracy	81.5%	79.6%	88.7%

5. Discussion

It could be shown that an expert system using diagnostic criteria published in the medical literature can perform successfully. As the ARA criteria were developed for classification purposes to get more uniform cohorts of RA patients for various clinical studies, an improvement might still be reached adding subjective clinical experience of a medical specialist for diagnostic purposes.

Acknowledgements. We thank I. Gröger for secretarial assistance and C. Schuh and F. Fischler for extended programming work. This research was partly supported by IBM Österreich.

References

[1] Adlassnig, K.-P. (1986) Fuzzy Set Theory in Medical Diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics* SMC-16, 260–265.

[2] Adlassnig, K.-P., Kolarz, G., Scheithauer, W., Effenberger, H. & Grabner, G. (1985) CADIAG: Approaches to Computer-Assisted Medical Diagnosis. *Computers in Biology and Medicine* 15, 315–335.

[3] Adlassnig, K.-P., Kolarz, G., Scheithauer, W. & Grabner, H. (1986) Approach to a Hospital-Based Application of the Medical Expert System CADIAG-2. *Medical Informatics* 11, 205–223.

[4] Arnett, F. C., Edworthy, St. M., Bloch, D. A., McShane, D. J., Fries, J. F., Cooper, N. S., Healey, L. A., Kaplan, St. R., Liang, M. H., Luthra, H. S., Medsger, Jr., Th. A., Mitchell, D. A., Neustadt, D. A., Pinals, R. S., Schaller, J. G., Sharp, J. T., Wilder, R. L. & Hunder, G. G. (1988) The American Rheumatism Association 1987 Revised Criteria for the Classification of Rheumatoid Arthritis. *Arthritis and Rheumatism* 33, 315–324.

[5] Kolarz, G. & Adlassnig, K.-P. (1986) Problems in Establishing the Medical Expert Systems CADIAG-1 and CADIAG-2 in Rheumatology. *Journal of Medical Systems* 10, 395–405.

[6] Ropes, M. W., Bennett, G. A., Cobb, S., Jacox, R. & Jassar, R. A. (1958) Revision of Diagnostic Criteria for Rheumatoid Arthritis. *Bulletin on Rheumatic Diseases* 9, 175–177.

[7] Steinbrocker, O., Traeger, C.H., & Batterman, R.C. (1949) Therapeutic Criteria in Rheumatoid Arthritis. *The Journal of the American Medical Association* 140, 659–662.