Strengths and Limitations of Automatic Knowledge Acquisition for the Medical Consultation System CADIAG-II/RHEUMA

Harald Leitich[1,2], Klaus-Peter Adlassnig[1], and Gernot Kolarz[3,4]

[1]Department of Medical Computer Sciences, Section on Medical Expert and Knowledge-Based Systems
University of Vienna Medical School, Spitalgasse 23, A-1090 Vienna, Austria
e-mail: harald.leitich@akh-wien.ac.at
[2]Department of Obstetrics and Gynecology
University of Vienna Medical School, Währinger Gürtel 18-20, A-1090 Vienna, Austria
[3]Clinic for Rheumatic Disease of the Social Insurance Company for Trade and Industry
Adolfine Malchergasse 1, A-2500 Baden, Austria
[4]Institute for Rheumatology
Marchetstraße 78, A-2500 Baden, Austria

Abstract. *To find out if semiautomatic knowledge acquisition may help to define or refine symptom-diagnosis relationships in CADIAG-II/RHEUMA's knowledge base, an analysis of the statistical relationships between selected CADIAG-II/RHEUMA symptoms and the diagnosis of rheumatoid arthritis (RA) was carried out using a large database of rheumatological patient charts. This study showed that both size and composition of patient and control groups must be carefully chosen before the results of the statistical analysis may serve as a basis to define or refine symptom-diagnosis relationships in the knowledge base of CADIAG-II/RHEUMA.*

1. Introduction

CADIAG-II is a consultation system for internal medicine which is based on fuzzy set theory and fuzzy logic and which was developed at the Department of Medical Computer Sciences of the University of Vienna [1–7]. In this system, symptoms and diagnoses are formalized as fuzzy sets, which are characterized by fuzzy membership functions [7,8]. In a given patient symptoms or diagnoses can thus be definitely present ($\mu=1$), partially present ($0<\mu<1$) or definitely absent ($\mu=0$). Relationships between symptoms and diagnoses are characterized by two aspects: (1) the frequency of occurrence degree and (2) the strength of confirmation degree both of which also take fuzzy values in [0,1].

CADIAG-II's rheumatological knowledge base was first developed by Kolarz as a knowledge base for the system CADIAG-I and was later modified and expanded for the CADIAG-II system [9]. It currently contains 1126 symptoms and 170 documented diagnoses. In order to evaluate the performance of CADIAG-II/RHEUMA, a large database of patient charts from a 140-bed rheumatological hospital in Baden/Austria has been built up subsequently, currently containing data of more than 3500 patients with a large variety of rheumatic conditions.

Semiautomatic acquisition of rheumatological knowledge for CADIAG-II has been studied before [6]. Using 2 x 2 tables (listing true positive, false positive, true negative, and false negative results) to analyze the statistical relationship between a symptom S and a diagnosis D, the frequency of occurrence degree can be statistically interpreted as P(S/D) or the rate of sensitivity and the strength of confirmation degree as P(D/S) or the positive predictive value. Thus, calculations of P(S/D) and P(D/S) might serve as a statistical basis to define or refine symptom-diagnosis relationships in CADIAG-II's knowledge base. A reformulation of these relationships as relative sigma-counts ΣCount(S/D) for the frequency of occurrence and ΣCount(D/S) for the strength of confirmation was done in [2].

In the present study we focused on the diagnosis of rheumatoid arthritis and we wanted to find out if semiautomatic knowledge acquisition would be helpful to verify or revise CADIAG-II's knowledge about this disease.

2. Methods

With the computerized records of 154 patients with rheumatoid arthritis (RA) and 3098 control patients with other rheumatological diagnoses, 2 x 2 tables, sensitivity and specificity rates, as well as positive predictive values (PPV) were consecutively calculated to show the statistical relationships between each CADIAG-II symptom and the diagnosis of RA. Symptoms were assumed to be either present, absent, or unknown in a given patient. For reasons of simplicity, we will present the results obtained for only a small set of symptoms that are based on the 1987 revised criteria for the classification of RA, published by the American College of Rheumatology (ACR) [10]. To demonstrate how the positive predictive value of a symptom for a diagnosis is strongly influenced by the prevalence of the diagnosis in the study population, we also calculated the normalized positive predictive value (normalized PPV) for each symptom by correcting for the different sizes of RA and control groups. Finally, we compared the results obtained with the CADIAG-II patient database with the results published by the ACR which were obtained with 262 RA patients and 262 control patients [10].

3. Results

In Tables 1 and 2, numbers of RA and control patients, sensitivity and specificity rates, and positive predictive values are displayed both for our study population and the ACR study. As shown from the results of the ACR study, all symptoms included had a strong statistical relationship to the diagnosis of RA. In contrast, the results obtained with the CADIAG-II database tended to show lower sensitivity rates and positive predictive values and higher specificity rates.

Symptom	RA patients (*N*)	Control patients (*N*)	Sensitivity (%)	Specificity (%)	PPV (%)	Normalized PPV (%)
Morning stiffness (> 1 hour)	148	3070	14.9	96.4	16.7	80.6
Swelling of 3 or more joint areas	154	3098	70.8	95.4	43.1	93.8
Swelling of the wrist, MCP or PIP joints	154	3098	85.1	91.6	33.4	91.0
Symmetric joint swelling	154	3098	81.2	90.6	30.0	89.6
Rheumatoid nodules	154	3098	13.0	98.3	27.0	88.2
Rheumatoid factor positive	152	3051	37.5	98.6	57.6	96.5
Radiographic changes typical of RA	137	84	100.0	0.0	62.0	50.0
4 out of 7 criteria positive	146	3000	72.6	98.7	72.6	98.2

Table 1: Results from CADIAG-II/RHEUMA patient database.

Symptom	RA patients (*N*)	Control patients (*N*)	Sensitivity (%)	Specificity (%)	PPV (%)
Morning stiffness (> 1 hour)	255	254	81.1	57.3	65.7
Swelling of 3 or more joint areas	254	253	90.7	84.0	85.2
Swelling of the wrist, MCP or PIP joints	262	261	96.6	74.8	79.3
Symmetric joint swelling	262	261	94.3	74.3	78.7
Rheumatoid nodules	260	259	43.4	97.7	95.0
Rheumatoid factor positive	250	207	80.4	87.0	88.2
Radiographic changes typical of RA	220	190	77.2	93.7	93.4
4 out of 7 criteria positive	262	262	91.2	89.3	89.5

Table 2: Results from ACR study [10].

4. Discussion

This study shows that the results from a statistical analysis of a patient database can only be a first step in the definition of symptom-diagnosis relationships. The differences between the results in the present and the ACR study demand a more detailed analysis of the underlying differences in study populations before statistical results may serve as a basis to define symptom-diagnosis relationships.

In the ACR study that primarily intents to find criteria for the homogenous classification of RA patients for clinical trials, only patients with a definite diagnosis of RA were included in the RA group whereas in our database a large percentage of RA patients tended to be at an early disease stage with less pronounced disease features. The control group in the ACR study included a much larger percentage of patients with other inflammatory rheumatic disorders compared to the control group of our database in which the majority of patients was affected by degenerative rheumatic disorders. For some symptoms, as a consequence, sensitivity rates in our population tended to be lower and specificity rates were higher. Positive predictive values were especially low in our population because of the different sizes of the RA and control groups and normalization of the positive predictive value was helpful to eliminate this bias. Thus, the composition of both study and control

groups must be carefully chosen before the results of the statistical analysis may be used to define or refine symptom-diagnosis relationships.

Because only positive X-ray signs were documented in our patient database, the results of the statistical analysis of the symptom "positive radiographic changes typical of RA" cannot be used to define its relationship to RA. For this and similar symptoms it would be necessary to conclude that symptoms are definitely absent if the respective examination has been carried out and no positive sign was recorded, a feature that is not included in the present version of the knowledge acquisition program. This strategy may be extended to define certain unknown symptoms such as biopsy results as being definitely absent, because the respective invasive examinations would certainly have been carried out if there were a chance that they would show a positive test result.

Other additional features of the knowledge acquisition program, which have already been planned as part of the ongoing MedFrame/CADIAG-IV project include the interpretation of the positive associations between symptoms and diagnoses (frequency of occurrence of S with D, strength of confirmation of S for D) as well as possible negative associations (frequency of occurrence of S with ¬D, strength of exclusion of S for ¬D) as relative sigma-counts [11,12].

5. References

[1] Adlassnig, K.-P. (1980) A Fuzzy Logical Model of Computer-Assisted Medical Diagnosis. *Methods of Information in Medicine* 19, 141–148.

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

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

[4] Adlassnig, K.-P., Kolarz, G., and Scheithauer, W. (1985) Present State of the Medical Expert System CADIAG-2. *Methods of Information in Medicine* 24, 13–20.

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

[6] Adlassnig, K.-P. and Kolarz, G. (1986) Representation and Semiautomatic Acquisition of Medical Knowledge in CADIAG-1 and CADIAG-2. *Computers and Biomedical Research* 19, 63–79.

[7] Adlassnig, K.-P. (1988) Uniform Representation of Vagueness and Imprecision in Patient's Medical Findings Using Fuzzy Sets. In Trappl, R. (ed.) *Cybernetics and Systems´88*, Kluwer Academic Publishers, Dordrecht, 685–692.

[8] Zadeh, L.A. (1965) Fuzzy Sets. *Information and Control* 8, 338–353.

[9] Adlassnig, K.-P. and Kolarz, G. (1982) CADIAG-2: Computer-Assisted Medical Diagnosis Using Fuzzy Subsets. In Gupta, M.M. and Sanchez, E. (eds.) *Approximate Reasoning in Decision Analysis.* North-Holland Publishing Company, Amsterdam, 219–247.

[10] Arnett, F.C., Edworthy, S.M., Bloch, D.A. et al. (1988) The American Rheumatism Association 1987 Revised Criteria for the Classification of Rheumatoid Arthritis. *Arthritis and Rheumatism* 33, 315–324.

[11] Rothenfluh, T.E., Boegl, K., and Adlassnig, K.-P. (2000) Representation and Acquisition of Knowledge for a Fuzzy Medical Consultation System. In Szczepaniak, P.S., Lisboa, P.J., and Tsumoto, S. (eds.) *Fuzzy Systems in Medicine.* Springer-Verlag, Heidelberg, 636–651.

[12] Leitich, H. (1995) Anforderungen an ein Wissenserwerbssystem für das Medizinische Expertensystem CADIAG-IV. *Diplomarbeit*. Technische Universität Wien, Österreich.