Semi-Automatic Knowledge Acquisition of Fuzzy Sets and Fuzzy Relationships on the Example of Laboratory Test Results with Different Types of Hepatitis

M. Schuerz [1], K.-P. Adlassnig [1], C. Lagor [2], B. Schneider [3], G. Grabner [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: {Michael.Schuerz | kpa}@akh-wien.ac.at

[2] Department of Medical Informatics, The University of Utah, Salt Lake City, USA
e-mail: Charles.Lagor@m.cc.utah.edu

[3] Department of Medical Statistics and Documentation, University of Vienna Medical School, Schwarzspanierstrasse 17, A-1090 Vienna, Austria, e-mail: Barbara.Schneider@univie.ac.at

[4] Professor emeritus of the Second Department for Gastroenterology and Hepatology and of the Department of Medical Computer Sciences, University of Vienna Medical School
e-mail: Georg.Grabner@teleweb.at

Background

Knowledge acquisition and representation are one of the central challenges for the successful construction and use of medical expert and knowledge-based systems in clinical practice. Since medical knowledge is immanently vague over wide ranges, fuzzy sets are used to deal with uncertain linguistic medical concepts such as *reduced*, *normal*, *elevated*, and *highly elevated*. During the diagnostic and therapeutic process observed symptoms, signs, and test results are converted into fuzzy compatibility values reaching from zero to unity with the linguistic medical concepts under consideration in a *data-to-symbol conversion* component of a medical expert system. Then these "fuzzified" findings are applied to establish confirmed and excluded diagnostic hypotheses by means of a knowledge base, which contains fuzzy relationships—generated by fuzzy relations—for the *frequency of occurrence* of findings with diseases and the *strength of confirmation* of findings for diseases.

Objective

The aim of this study was to construct fuzzy sets in the form of *S*- and π-shaped fuzzy membership functions semi-automatically from data of one reference group of healthy individuals and six disease groups of patients suffering from liver diseases for the observed laboratory parameters for the linguistic medical concepts *reduced*, *normal*, *elevated*, and *highly elevated.* After this, fuzzy relations were formed between linguistic medical concepts and diagnoses in the field of hepatitis diseases [1].

Material and Methods

Laboratory test results from sample patients from reference groups of healthy individuals and from the following hepatitis groups were investigated: type A hepatitis, type B hepatitis, type C hepatitis, chronic hepatitis, alcoholic hepatitis, and hepatitis B carriers with values for the laboratory parameters bilirubin, alanine aminotransferase, aspartate aminotransferase gamma-glutamyltranspeptidase, alkaline phosphatase, lactate dehydrogenase, and for the electrophoresis laboratory parameters albumin, alpha 1 globulin, alpha 2 globulin, beta globulin, and gamma globulin. Considering the reference groups of the previously mentioned laboratory parameters fuzzy membership functions were constructed for the linguistic medical concept *normal* with parameterless distribution-free statistical methods. By this the

minimum, the 5%-percentile, the 95%-percentile, and the maximum were calculated from data and used as parameters for *S*- and π-shaped fuzzy membership functions [2].
In the next step fuzzy membership functions for *reduced* and *elevated* were formed. Since these concepts are complementary to the concept *normal*, the corresponding *S*-membership functions of *normal* were reflected on their inflection points to obtain the fuzzy membership functions of *reduced* and *elevated*. Finally a differentiation of a subset *highly elevated* from *elevated* could be obtained. An iterative gradient method using ordered sets of maxima and 95%-percentiles from the diseases and the reference groups was applied to find a threshold between *elevated* and *highly elevated*. From this, fuzzy relations between uncertain findings and diagnoses were generated for the frequency of occurrence of findings with diseases and the strength of confirmation of symptoms for diseases using relative sigma-counts [3]. In this case these relative sigma-counts are the conditional probabilities $P(D_j/S_i)$ and $P(S_i/D_j)$ with uncertain (fuzzy) events S_i and D_j with S_i representing the fuzzy compatibility value of a finding of a specific patient and D_j representing the corresponding fuzzy compatibility value of a diagnosis of a specific patient.

Results

Based on laboratory test results, 41 fuzzy sets for the representation of the linguistic medical concepts *reduced*, *normal*, *elevated*, and *highly elevated* were obtained for the test results of the eleven above-mentioned laboratory and electrophoresis parameters (Fig. 1).

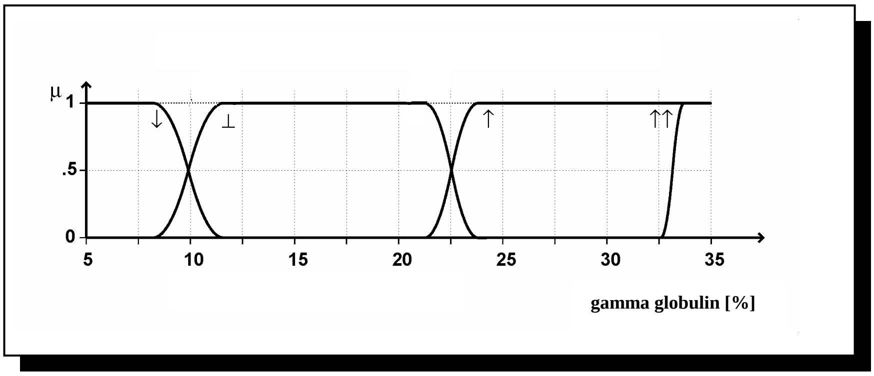

Fig. 1: Fuzzy membership functions for *reduced* ($\mu_{\downarrow}$), *normal* ($\mu_{\perp}$), *elevated* ($\mu_{\uparrow}$), and *highly elevated* ($\mu_{\uparrow\uparrow}$) for gamma globulin.

From this, 574 fuzzy relations for the generation of the frequency of occurrence of the linguistic medical concepts *reduced*, *normal*, *elevated*, and *highly elevated* for every above-mentioned laboratory parameter with the hepatitis and reference groups as well as the strength of confirmation of the previously mentioned findings for the hepatitis and reference groups were calculated (to be applied during the inference process of medical expert systems) (Tab. 1). Hereby fuzzy relations with values equal zero as well as those equal one need to be checked by a physician according to their frequency of occurrence respectively strength of confirmation in the parent population since these values are calculated from sample data and hence may contain biases.

	type A hepatitis		type B hepatitis		reference group	
	μ_O	μ_C	μ_O	μ_C	μ_O	μ_C
gamma globulin						
reduced	0.00	0.00	0.01	0.00	0.02	0.48
normal	0.40	0.03	0.64	0.00	0.96	0.34
elevated	0.60	0.12	0.35	0.00	0.02	0.02
highly elevated	0.00	0.00	0.00	0.00	0.00	0.00

Tab. 1: Fuzzy relations between the uncertain symptoms *reduced* ($\mu_{\downarrow}$), *normal* ($\mu_{\perp}$), *elevated* ($\mu_{\uparrow}$), and *highly elevated* ($\mu_{\uparrow\uparrow}$) for gamma globulin with type A hepatitis, type B hepatitis, and the reference group. The fuzzy relation μ_O indicates the frequency of occurrence of findings with diseases and μ_C indicates the strength of confirmation of findings for diseases.

Technical Specification

Laboratory test results and diagnoses were extracted from the data base of the hospital information system WAMIS [2] that runs on a IBM 2003 under VSE/ESA with CICS and transferred to Microsoft Excel 97 for calculation of the corresponding parameters for the *S*- and π-shaped fuzzy membership functions as well as to calculate fuzzy relations. Turbo Pascal 6.0 was used to write programs for the optimization of fuzzy sets parameters. Finally, the fuzzy membership functions were plotted in postscript with GNU-Plot MS-Windows version 3.5.

Conclusion

The aim of this study was to develop a feasible method of semi-automatic knowledge acquisition of fuzzy sets for linguistic medical concepts and fuzzy relations between medical findings and diagnoses. This could be shown with the proposed method on the applied material. This method will remain semi-automatic because the results have to be checked for plausibility and—should the occasion arise—adapted by a physician for the intended application. Future work will cover not only an evaluation of different approaches concerning the semi-automatic construction of fuzzy sets and fuzzy relations but also an application of this method to different medical areas.

References

[1] Schürz, M., Adlassnig, K.-P., Lagor, C., Schneider, B. & Grabner, G. (1998) Halbautomatischer Wissenserwerb von Fuzzy-Mengen und Fuzzy-Relationen bei Hepatitiden*. Abstractband zum 2. Symposium "Medizinische Experten- und Wissensbasierte Systeme und Computergestützte OP-Navigation am AKH-Wien",* Universität Wien, 44. (see also: http://www.ai.univie.ac.at/imkai/wbsakh98/tofc.html).

[2] Schuerz, M., Adlassnig, K.-P., Lagor, C., Schneider, B. & Grabner, G. (1998) Supervised Construction of Fuzzy Sets from Sample Data: Laboratory Test Results Describing Different Forms of Hepatitis*. Workshop on Applications of Fuzzy Logic,* ERUDIT, 23—30.

[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] Grabner, G. (1985) (Hrsg.) *WAMIS – Wiener Allgemeines Medizinisches Informations-System,* Springer-Verlag, Berlin.