Annals of Nuclear Medicine Vol. 11, No. 2, 75–80, 1997

Diagnosis of chronic liver disease from liver scintiscans by artificial neural networks

Susumu SHIOMI,* Tetsuo KUROKI,* Maki KURIYAMA,* Hiroyasu MORIKAWA,* Kyoko MASAKI,* Naoko IKEOKA,* Takashi TANAKA,** Hozumi IKEDA*** and Hironobu OCHI***

*Third Department of Internal Medicine, **Department of Public Health and ***Division of Nuclear Medicine, Osaka City University Medical School

Artificial neural networks were used in the diagnosis of chronic liver disease based on liver scintiscanning. One hundred and thirty-seven patients with chronic liver disease (12 with chronic persistent hepatitis, 39 with chronic aggressive hepatitis, and 86 with cirrhosis) and 25 healthy controls were studied. Sixty-five subjects (10 healthy controls, 20 patients with chronic hepatitis, and 35 patients with cirrhosis of the liver) were used in the establishment of a neural network. Liver scintiscans were taken starting 20 min after the intravenous injection of 111 MBq of Tc-99m-phytate. The neural network was used to evaluate five items judged from information on liver scintiscans: the ratio of the sizes of the left and right lobes, splenomegaly, radioactivity in the bone marrow, deformity of the liver and distribution of radioactivity in the liver. The neural network was designed to distinguish between three liver conditions (healthy liver, chronic hepatitis and cirrhosis) on the basis of these five items. The diagnostic accuracy with the neural network was 86% for patients with chronic hepatitis and 93% for patients with cirrhosis. With conventional scoring, the accuracy was 77% for patients with chronic hepatitis and 87% for patients with cirrhosis. Our findings suggest that artificial neural networks may be useful for the diagnosis of chronic liver diseases from liver scintiscans.

Key words: chronic hepatitis; cirrhosis; liver scintiscan; artificial neural network; technetium-99m-phytate

INTRODUCTION

LIVER SCINTISCANS have been used for the diagnosis of both localized liver diseases[1–3] and diffuse hepatocellular diseases.[4–8] For the diffuse diseases, in particular, scintiscanning has greater diagnostic usefulness than other imaging methods such as computed tomography and ultrasonography,[9] but diagnosis from liver scintiscans must be done through subjective evaluation of the image. Fuzzy reasoning[10] and neural networks[11] are useful for evaluations that involve subjectivity. The use of artificial neural networks was described by Rumelhart et al.[11] in 1986. Such networks are computational models that simulate the function of nerve networks. The models have been applied to various fields[12] including medicine.[13–18] We examined the usefulness of a neural network in the evaluation of liver scintiscans for the diagnosis of diffuse hepatocellular disease.

Received April 1, 1996, revision accepted December 19, 1996.

For reprint contact: Susumu Shiomi, M.D., Third Department of Internal Medicine, Osaka City University Medical School, 1–5–7 Asahimachi, Abeno-ku, Osaka 545, JAPAN.

MATERIALS AND METHODS

Subjects

Liver scintiscanning was done in 162 subjects: 137 patients with hepatic viral infection and 25 control subjects found to have healthy livers when examined for suspected disease. Diagnoses from scintiscans were compared with the definitive diagnosis of hepatitis or cirrhosis based on histological examination of liver specimens obtained by laparoscopy or needle biopsy done under ultrasonic guidance. Results of the histological examinations, done in accordance with international established criteria,[19] showed

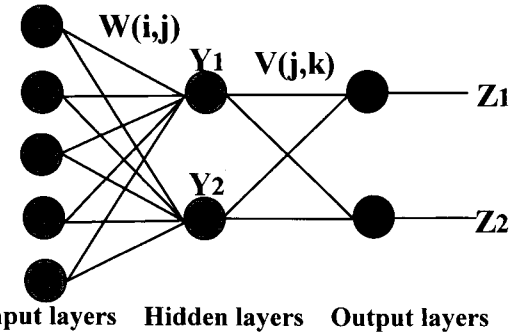

Fig. 1 Basic structure of the artificial neural network. A simple three-layer network is shown with five input nodes (left), two hidden layer nodes (center), and two output layers (right).

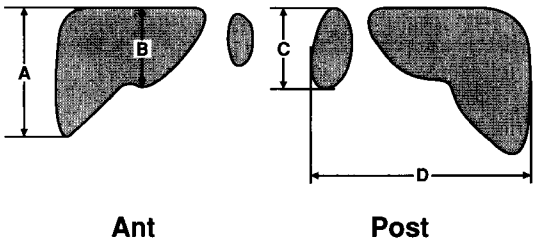

Fig. 2 Calculation of the left-to-right ratio of the liver and splenomegaly from scintiscans.

that 12 patients had chronic persistent hepatitis, 39 had chronic aggressive hepatitis, and 86 had cirrhosis. The scintiscans of 10 healthy controls, 20 patients with chronic hepatitis (3 patients with chronic persistent hepatitis and 17 with chronic aggressive hepatitis) and 35 patients with cirrhosis were chosen randomly from the scintiscans of these 162 subjects and used for establishment of a neural network. The remaining 15 controls and 82 patients with chronic liver diseases were used in testing the neural network.

Liver scintiscan

Liver scintiscans were taken starting 20 min after the intravenous injection of 111 MBq of Tc-99m-phytate. Images (400,000 counts for each) were obtained in anterior and posterior views with a scintillation camera (Siemens ZLC).

Artificial neural network

The principles of neural networks can be summarized as follows. Input data, which are represented by numbers between 0 and 1, are supplied to input units of the neural network. Then output data are provided from output units after two successive nonlinear calculations in hidden and output layers.

The artificial neural network we used has the structure shown in Fig. 1. We set the input layer at 5 units, the hidden layer at 2 units, and the output layer at 2 units. Two basic processes, learning and testing, are involved in a neural network. The neural network 'learns' from a back-propagation algorithm that has pairs of training input data and desired output data.[20] The internal parameters of the neural network are adjusted to minimize the difference between the actual output of the neural network and the desired output.

Learning by the neural network

Input data were used to evaluate five items: the ratio of sizes of the left and right lobes (X_1), splenomegaly (X_2) as seen on the scintigram, radioactivity in the bone marrow (X_3), deformity of the liver (X_4), and the distribution of radioactivity in the liver (X_5). To calculate the left-to-right ratio and splenomegaly, we measured the dimensions of the images of scintiscans as shown in Fig. 2. To use the values of these two parameters as the input signals for the neural network, we converted them as follows so that they might be in the range of 0 to 1[11]:

$$Z_{new} = \mathbf{a}Z_{old} + \mathbf{b} \qquad \text{Eq. 1}$$

where Z_{old} is the value before conversion and Z_{new} is the value after conversion, with **a** and **b** converted so that each parameter would have a minimum value of 0 and a maximum value of 1. For X_1, **a** = 1.14 and **b** = 0.42, and for X_2, **a** = 1.52 and **b** = 0.20. For radioactivity in the bone marrow, deformity of the liver, and the distribution of radioactivity in the liver, the value range of 0 to 1 was divided into five possible grades: 0, 0.25, 0.5, 0.75 and 1.0. If in posterior views the bone marrow was not visible (not radioactive), the grade given was 0; if the bone marrow was clearly visible, the grade was 0.5; if the bone marrow and the ribs were clearly visible, the grade was 1.0. In scoring for deformity, a liver without any deformity in the anterior view was given a score of 0; if the right and left lobes were seen to be indented at their interface, the score was 0.5; if in addition to the indentation, the right lobe was atrophic, the score was 1.0. In scoring radioactivity, a uniform distribution in the anterior view was graded 0, and a very uneven distribution was graded as 1.0.

We first used input data from 10 healthy controls (H_1 to H_{10}), 20 patients with chronic hepatitis (C_1 to C_{20}) and 35 patients with cirrhosis of the liver (L_1 to L_{35}). For the controls, X_1 was 0.12 to 0.23; X_2 was 0.02 to 0.32; X_3 was 0; X_4 was 0 to 0.25 and X_5 was 0. For the patients with chronic hepatitis, X_1 was 0.08 to 0.39; X_2 was 0.17 to 0.77; X_3 was 0 to 0.25; X_4 was 0 to 0.5 and X_5 was 0 to 0.25. For cirrhotic patients, X_1 was 0.06 to 0.72; X_2 was 0.17 to 0.82; X_3 was 0.5 to 1.0; X_4 was 0.25 to 1.0 and X_5 was 0 to 1.0. Learning was then started with the teaching signal of (0, 0) for a healthy liver, (1, 0) for chronic hepatitis and (1, 1) for cirrhosis.

Testing of the neural network

Finally, diagnosis was done with the neural network, with results for 97 subjects (15 healthy controls and 82 patients with chronic liver diseases) being analyzed.

Table 1 Conventional scoring of scintigrams

Score	Left/right ratio	Splenomegaly	Radioactivity in bone marrow	Deformity of the liver	Distribution of radioactivity
0	$B/A < 0.56$	$C/D < 0.26$	None	None	Uniform
1	$0.56 \leq B/A < 0.70$	$0.26 \leq C/D < 0.40$	Moderate	Moderate	Intermediate
2	$0.70 \leq B/A$	$0.40 \leq C/D$	Marked	Marked	Not uniform

Intermediate, between uniform and nonuniform. See Figure 2 for *A–D*.

Table 2 Output data of the neural network

Subjects	n	Hidden layer		Output layer	
		Y_1	Y_2	Z_1	Z_2
Healthy	10	0.02 (0, 0.12) [0 – 0.29]	0.96 (0.96, 0.97) [0.92 – 0.98]	0 (0, 0.01) [0 – 0.09]	0 (0, 0) [0 – 0]
Chronic hepatitis	20	0.98 (0.93, 1.00) [0.71 – 1.00]	0.94 (0.90, 0.96) [0.72 – 0.96]	1.00 (0.99, 1.00) [0.96 – 1.00]	0 (0, 0) [0 – 0.03]
Cirrhosis	35	1.00 (1.00, 1.00) [0.99 – 1.00]	0 (0, 0.02) [0 – 0.36]	1.00 (1.00, 1.00) [1.00 – 1.00]	1.00 (1.00, 1.00) [0.97 – 1.00]

Medians (25th, 75th percentiles). [Range of data is in square brackets]

Table 3 Results of learning by artificial neural network

Items		X_1, Left/right ratio	X_2, Splenomegaly	X_3, Radioactivity in bone marrow	X_4, Deformity	X_5, Distribution	Bias unit
Hidden layer	$W(i, 1)$	0.07	6.93	1.91	3.34	0.32	–2.5
	$W(i, 2)$	0.77	–0.42	–2.49	–0.51	0.01	0.63

Conventional scoring of scintigrams

Scoring by a conventional method was done on a three-point scale for the same five items (Table 1[21]). The sum of the scores for each patient was used as the scintiscore. Neural network results were compared with those of the conventional method.

RESULTS

Learning by means of the neural network

The relationship between error in the learning process by means of the neural network and the diagnostic accuracy was examined. When 1,000 learning steps were accomplished, the error was 0.0001, and the diagnostic accuracy was 100%. When the number of learning steps was increased further, the decrease in error per unit time became smaller.

Table 2 shows the values for the hidden and output layers obtained through learning. The output of the hidden layer was expressed in the equation $f(x) = 1/[1 + \exp(-x)]$, and the output layer was expressed in the linear function $f(x) = x$. The hidden layer Y_1 was almost 0 for the healthy livers, and 1 or nearly 1 for chronic hepatitis and cirrhosis. In addition, the hidden layer Y_2 was almost 1 for the normal liver and chronic hepatitis, and close to 0 for cirrhosis. The output layers Z_1 and Z_2 were close to 0 and 0, respectively, for the healthy livers, close to 1 and 0 for chronic hepatitis, and close to 1 and 1 for cirrhosis; in short, the output layers were close to the teaching signals.

Table 3 shows the weighting factors and offset values corresponding to X_1 to X_5 from liver scintigrams, after learning by means of the artificial neural network. The weighting factors for the hidden layer Y_1 were $W(i, 1)$ and those for the hidden layer Y_2 were $W(i, 2)$.

Testing of the neural network

Patient 1. Chronic aggressive hepatitis was diagnosed histologically in a 51-year-old man. From the scintiscan (Fig. 3), the left/right ratio (B/A) was 0.52, splenomegaly (C/D) was 0.33, the bone marrow radioactivity was 0.25, deformity of the liver was 0.25 and the distribution of radioactivity in the liver was 0.25. The diagnosis by means of the artificial neural network was chronic hepatitis ($Z_1 = 1.00$ and $Z_2 = 0.01$).

Patient 2. Cirrhosis was diagnosed histologically in a 63-year-old man. From the scintiscan (Fig. 4), the left/

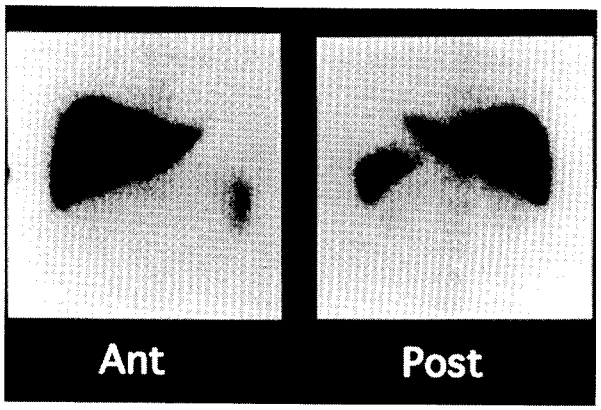

Fig. 3 Scintiscan, patient 1.

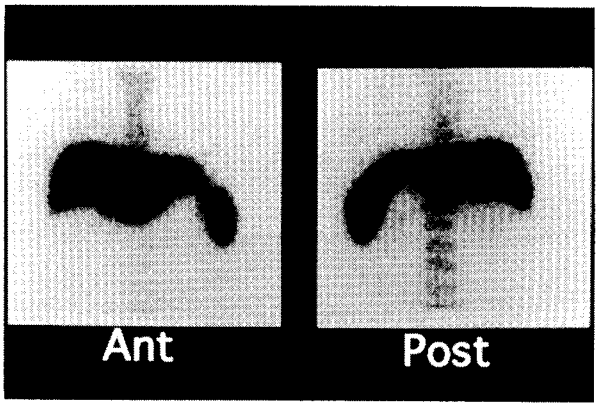

Fig. 4 Scintiscan, patient 2.

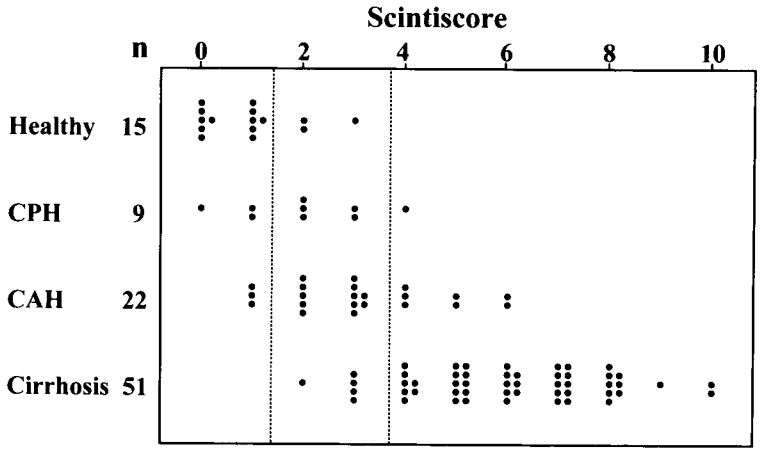

Fig. 5 Conventional scintiscores for the 97 subjects used in testing of the neural network.

Table 4 Correlation between the diagnosis by the artificial neural network and the histological diagnosis

	Diagnosis	Neural network		
		Normal	Chronic hepatitis	Cirrhosis
Histological finding	Healthy	12	3	0
	CPH	3	6	0
	CAH	1	18	3
	Cirrhosis	0	4	47

CPH, chronic persistent hepatitis; CAH, chronic aggressive hepatitis.

right ratio was 0.96, splenomegaly was 0.57, bone marrow radioactivity was 0.75, deformity of the liver was 0.75, and the distribution of radioactivity in the liver was 0.50. The diagnosis by means of the artificial neural network was cirrhosis of the liver ($Z_1 = 1.00$ and $Z_2 = 0.99$).

Table 4 shows the correlation between the diagnosis by means of the artificial neural network and the diagnosis based on histological findings. The results of conventional scoring for the same subjects are given in Fig. 5. Dotted lines were drawn at the positions that gave the greatest accuracy based on the histological diagnoses. Table 5 compares the results obtained with the neural network to those obtained with conventional scoring.

Table 5
Sensitivity, specificity, and accuracy for chronic hepatitis

	Sensitivity	Specificity	Accuracy
Neural network	24/31 (77%)	59/66 (89%)	83/97 (86%)
Conventional scoring	17/31 (55%)	58/66 (88%)	75/97 (77%)

Sensitivity, specificity, and accuracy for cirrhosis of the liver

	Sensitivity	Specificity	Accuracy
Neural network	47/51 (94%)	43/46 (91%)	90/97 (93%)
Conventional scoring	46/51 (90%)	38/46 (83%)	84/97 (87%)

Numbers of patients are shown, with percentages in parentheses.

DISCUSSION

The neural network is useful for letter recognition, voice recognition, and mechanical translation, all operations that require subjective judgment through sophisticated parallel distributed processing, and that cannot be achieved by computers of the Von-Neumann type. Diagnosis by liver scintigraphy is a possible application of neural networks because the diagnosis involves subjective judgment.

The accuracy of diagnosis of cirrhosis with the neural network was 93%, higher than has been reported so far for scanning techniques.[9] After training of the neural network, the hidden layer Y_1 was close to 0 for the healthy controls, and almost 1 for chronic hepatitis and cirrhosis. In other words, Y_1 could be used to discriminate between a healthy liver and one with chronic disease. Splenomegaly gave the largest absolute values for $W(i, 1)$, followed by deformity of the liver and radioactivity in bone marrow. These factors therefore contributed much to the discrimination between a healthy liver and one with chronic disease. Similarly, Y_2 was almost 1 for healthy livers and chronic hepatitis, and close to 0 for cirrhosis. Y_2 could therefore be used to judge whether the disease was cirrhosis of the liver or not. The radioactivity in bone marrow gave the largest absolute values for $W(i, 2)$, followed by left/right ratio and splenomegaly. These factors contributed much in judging whether the disease was cirrhosis of the liver or not. The items selected were left/right ratio, splenomegaly, radioactivity in bone marrow and deformity of the liver, because they contributed most of the items investigated. The distribution of radioactivity in the liver was because omission of this distribution violated the condition we set for teaching signals showing that a healthy liver was (0, 0).

We have tried fuzzy reasoning for the diagnosis of chronic liver disease from liver scintiscans,[22] obtaining a sensitivity of 76%, specificity of 91% and accuracy of 86% for chronic hepatitis, and a sensitivity of 91%, specificity of 95% and accuracy of 93% for cirrhosis, values not much different from those reported here. Fuzzy reasoning and neural networks are useful in the identification of nonlinear input-output relationships, but problems remain to be solved if they are to be used. The neural network has a autotraining function which enables identification of a nonlinear input-output relationship even when the model environment is changing. The main cause of decreases in the generalization ability of neural networks is overtraining. To avoid overtraining, we ended neural network training when the error was 0.0001. Another method to prevent avoidable decreases in the generalization ability is to use only the minimum number of units and layers in the hidden layer. When the number of hidden layer units was three, the accuracy of the diagnosis of chronic hepatitis was 84% and that of cirrhosis was 91%. The percentages were higher with the two units we decided on. When two hidden layers were used, these values were 82% and 90%. The percentages were higher with one layer, which we therefore used, but the neural network cannot clearly express either the status or results of learning. The main disadvantage of fuzzy reasoning is that the fuzzy rule must be established arbitrarily, and results vary depending on how it is established. Neural networks have been combined with fuzzy reasoning in an attempt to make up for the defects of each method.[23] In that attempt, the fuzzy rule was established with the diagnostic function of a neural network, or else the diagnostic results from the neural network were expressed by fuzzy reasoning. Artificial neural networks may be useful for the diagnosis of chronic liver diseases from liver scintiscans.

REFERENCES

1. Yuhl ET, Stirrett LA, Cassen B. Use of colloidal Au-198 for obtaining scintigrams of the liver. *Nucleonics* 11: 54–56, 1953.
2. Lunia S, Parthasarathy KL, Bakshi S, Bender MA. An evaluation of Tc-99m-sulfur colloid liver scintiscans and their usefulness in metastatic workup: a review of 1,424 studies. *J Nucl Med* 16: 62–65, 1975.
3. Fritz SL, Preston DF, Gallagher JH. ROC analysis of diagnostic performance in liver scintigraphy. *J Nucl Med* 22: 121–128, 1981.
4. Wagner HN Jr, McAfee JG, Mozley JM. Diagnosis of liver disease by radioisotope scanning. *Arch Intern Med* 107: 324–334, 1961.
5. Whang KS, Fish MB, Pollycove M. Evaluation of hepatic photoscanning with radioactive colloidal gold. *J Nucl Med* 6: 494–505, 1965.
6. Geslien GE, Pinsky SM, Poth RK, Johnson MC. The sensitivity and specificity of Tc-99m-sulfur colloid liver imaging in diffuse hepatocellular disease. *Radiology* 118: 115–119, 1976.
7. Drum DE, Beard JO. Liver scintigraphic features associated with alcoholism. *J Nucl Med* 19: 154–160, 1978.
8. Picard D, Infante-Rivard C, Villeneuve J-P, Chartrand R, Picard M, Carrier L. Extrahepatic uptake of technetium-99m-phytate: a prognostic index in patients with cirrhosis. *J Nucl Med* 31: 436–440, 1990.
9. Meek DR, Mills PR, Gray HW, Duncan JG, Russell RI, McKillop JH. A comparison of computed tomography, ultrasound and scintigraphy in the diagnosis of alcoholic liver disease. *Br J Radiol* 57: 23–27, 1984.
10. Zadeh LA. Fuzzy sets. *Information and Control* 8: 338–353, 1965.
11. Rumelhart DE, Hilton GE, William RJ. Learning internal representations by error propagation. *In:* Rumelhart DE, McClelland JL, eds. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1.* Cambridge, MA: MIT Press, pp. 318–362, 1986.
12. Poggio T, Girosi F. Regularization algorithms for learning that are equivalent to multilayer networks. *Science* 247: 978–982, 1990.
13. Asada N, Doi K, MacMahon H, Montner SM, Giger ML,

Abe C, et al. Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung diseases: pilot study. *Radiology* 177: 857–860, 1990.
14. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. *J Nucl Med* 33: 272–276, 1992.
15. Scott JA, Palmer EL. Neural network analysis of ventilation-perfusion lung scans. *Radiology* 186: 661–664, 1993.
16. Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE. Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer. *Radiology* 187: 81–87, 1993.
17. Tourassi GD, Floyd CE, Sostman HD, Coleman RE. Acute pulmonary embolism: artificial neural network approach for diagnosis. *Radiology* 189: 555–558, 1993.
18. Porenta G, Dorffner G, Kundrat S, Petta P, Duit-Schedlmayer J, Sochor H. Automated interpretation of planar thallium-201-dipyridamole stress-redistribution scintigrams using artificial neural networks. *J Nucl Med* 35: 2041–2047, 1994.
19. Scheuer PJ, Thaler H. Acute and chronic hepatitis revisited: review by an international group. *Lancet* 2: 914–919, 1977.
20. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 323: 533–536, 1986.
21. Taniguchi S, Tanaka S, Okamura T, Sazaki A, Takemoto K, Ochi H, et al. Comparison of colloid liver scan, computed tomography and gray-scale ultrasonography in the evaluation of the diffuse liver disease proven by the peritoneoscopic biopsy [Abstract]. *KAKU IGAKU (Jpn J Nucl Med)* 22: 752, 1985.
22. Shiomi S, Kuroki T, Jomura H, Ueda T, Ikeoka N, Kobayashi K, et al. Diagnosis of chronic liver disease from liver scintiscans by fuzzy reasoning. *J Nucl Med* 36: 593–598, 1995.
23. Ogawa Y, Morita T, Kobayashi K. A fuzzy document retrieval system using the keyword connection matrix and a learning method. *Fuzzy Sets Systems* 39: 163–179, 1991.