Cardiac Risk Stratification in Renal Transplantation Using a Form of Artificial Intelligence

Thomas F. Heston, MD, Douglas J. Norman, MD, John M. Barry, MD, William M. Bennett, MD, and Richard A. Wilson, MD

The purpose of this study was to determine if an expert network, a form of artificial intelligence, could effectively stratify cardiac risk in candidates for renal transplant. Input into the expert network consisted of clinical risk factors and thallium-201 stress test data. Clinical risk factor screening alone identified 95 of 189 patients as high risk. These 95 patients underwent thallium-201 stress testing, and 53 had either reversible or fixed defects. The other 42 patients were classified as low risk. This algorithm made up the "expert system," and during the 4-year follow-up period had a sensitivity of 82%, specificity of 77%, and accuracy of 78%. An artificial neural network was added to the expert system, creating an expert network. Input into the neural network consisted of both clinical variables and thallium-201 stress test data. There were 5 hidden nodes and the output (end point) was cardiac death. The expert network increased the specificity of the expert system alone from 77% to 90% ($p < 0.001$), the accuracy from 78% to 89% ($p < 0.005$), and maintained the overall sensitivity at 88%. An expert network based on clinical risk factor screening and thallium-201 stress testing had an accuracy of 89% in predicting the 4-year cardiac mortality among 189 renal transplant candidates. *©1997 by Excerpta Medica, Inc.*

(Am J Cardiol 1997;79:415–417)

In the United States, registrations on the United Network of Organ Sharing national waiting list are more than three times the number of cadaver kidneys available for transplantation.[1] Organ life can be increased by transplanting kidneys into the healthiest transplant candidates. Cardiac disease is the most common cause of death among renal transplant recepients.[2] Thus, risk stratification before transplantation includes an assessment of cardiovascular status. This study attempts to use 2 complimentary forms of artificial intelligence in order to improve cardiac risk stratification. First, an "expert system"[3] based on deductive reasoning (applying results of large population studies to individual cases, e.g., a decision tree) was used. However, this lacked specificity, so an "artificial neural network"[4–6] was added to create a more specific "expert network"[7] using inductive reasoning[8] (i.e., making generalizations based on individual cases).

METHODS

Expert system design: Originally, the expert system was prospectively applied to the 189 consecutive patients being considered for renal transplantation at Oregon Health Sciences University between June 1987 and September 1988. Patients were followed from the time of entry up until August 1992. The median follow-up was 46.8 months (range 0.5 to 62) and 94% of patients were followed for at least 1 year. This algorithm is described in detail elsewhere.[9] The initial decision point on the algorithm was the presence or absence of ≥1 of the following 5 high risk factors: congestive heart failure, a history of angina, an abnormal electrocardiogram (except for left ventricular hypertrophy), insulin-dependent diabetes mellitus, or age ≥50 years. Patients without any of these 5 clinical variables were placed in the low-risk category and no further cardiac workup was undertaken. Those with ≥1 of these risk factors underwent thallium-201 stress testing with either exercise or dipyridamole. For this study, patients with normal perfusion were placed in the low-risk category. Patients with ≥1 fixed or reversible defect were placed in the high-risk category. This algorithm alone constituted the expert system.

Neural network architecture: Only patients placed in the high-risk category by the expert system were analyzed by the neural network. Inputs into the neural network included 15 clinical variables and 5 variables from the thallium-201 stress test (Table I), giving a total of 20 input nodes. The neural network had 5 hidden nodes in 1 layer. This layer of 5 hidden nodes is a layer of neurones that does not connect to the outside world but connects to the output node. There was 1 output node consisting of cardiac death. The artificial neural network was created using a commercial software program run on a desktop personal computer under Microsoft Windows (Brainmaker version 3.1, Nevada City, California). The artificial neural network was trained to convergence (at which point the network has moved toward a stable state after the input pattern has been applied) on $\frac{1}{2}$ of the high-risk patients, then blindly tested only once on the other $\frac{1}{2}$. Patients were randomly assigned

From the Division of Nuclear Medicine, Department of Radiology; Division of Nephrology and Division of Cardiology, Department of Medicine; and the Division of Urology, Department of Surgery, Oregon Health Sciences University, Portland, Oregon. Manuscript received May 6, 1996; revised manuscript received and accepted September 10, 1996.
Address for reprints: Richard A. Wilson, MD, Division of Cardiology, L-462, Oregon Health Sciences University, 3181 S.W. Sam Jackson Park Road, Portland, Oregon 97201.

0002-9149/97/$17.00
PII S0002-9149(96)00778-3

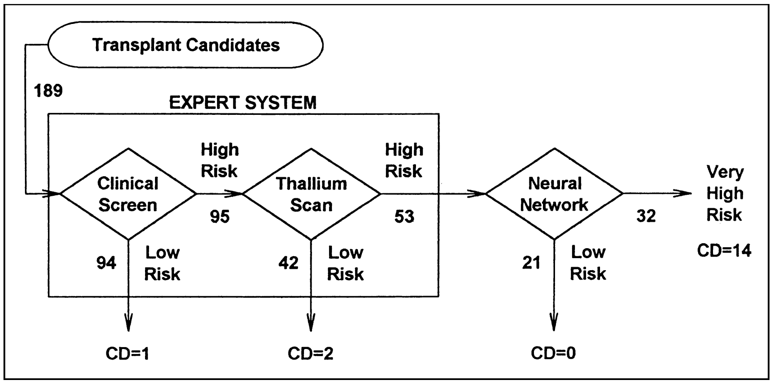

FIGURE 1. The addition of an artificial neural network to the expert system enabled an additional 21 patients to be considered at low risk for cardiac death (CD).

TABLE I Clinical and Thallium Stress Test Variables

Clinical variables
1. Age (<40, 40–49, 50–59, >60)
2. Gender
3. Insulin-dependent diabetes mellitus (present/absent)
4. Q waves on electrocardiogram (yes/no)
5. Angina (yes/no)
6. Congestive heart failure (yes/no)
7. Abnormal baseline electrocardiogram (yes/no)
8. Hypertension (yes/no)
9. Smoker (yes/no)
10. Previous coronary artery bypass graft (yes/no)
11. Taking digoxin (yes/no)
12. Taking calcium channel blockers (yes/no)
13. Taking β blockers (yes/no)
14. Taking transdermal nitroglycerin (yes/no)
15. Cholesterol >240 mg/dl (yes/no)
Thallium-201 stress test variables
1. ST-segment abnormality on electrocardiogram
2. Chest pain during stress test
3. Number of reversible defects
4. Number of fixed defects
5. Lung/heart ratio >0.50

TABLE II Components of Screening Algorithm

	Cardiac Death			
	Yes	No	Chi-Square	p Value
Clinical risk factors				
Present	16	79		
Absent	1	93	14.5	<0.001
Perfusion defect				
Present	14	39		
Absent	2	40	7.6	<0.01
Expert system (clinical + thallium)				
High risk	14	39		
Low risk	3	136	26.0	<0.001
Artificial intelligence				
High risk	7	9		
Low risk	0	11	5.0	<0.05
Expert network				
High risk	14	18		
Low risk	3	154	55.6	<0.001

to either the testing or training group, with the exception that $\frac{1}{2}$ of the cardiac deaths were intentionally placed in the training group, and the other $\frac{1}{2}$ in the testing group. Results of the blinded testing were applied to the entire high-risk cohort to estimate the incremental value of the neural network.

Thallium scintigraphy: Planar thallium imaging was performed in the anterior, 45°, and 70° left anterior oblique projections. Quantitative circumferential count profile analysis was performed as previously described.[9,10]

Statistical analysis: Yates' corrected chi-square values were calculated for the 2 × 2 contingency tables.[11] Z scores were calculated to compare changes in sensitivity, specificity, and accuracy made by the addition of the neural network to the expert system.[12]

RESULTS

Clinical risk factor screening: This clinical screening algorithm had a sensitivity, specificity, and accuracy of 94%, 54%, and 58%, respectively, for detecting cardiac death among the 189 renal transplant candidates (Table II).

Thallium-201 stress testing: Patients with ≥1 of the clinical high-risk factors were further stratified by thallium-201 stress testing. The sensitivity, specificity, and accuracy of the thallium-201 stress test on the 95 patients identified as high risk by the clinical screening algorithm were 88%, 51%, and 57%, respectively (Table II). The addition of thallium-201 stress testing to the clinical screening algorithm improved the overall accuracy from 58% to 78% ($p < 0.001$). Together, the clinical screening algorithm combined with thallium-201 stress testing constituted the expert system (Table II).

Artificial neural network: The 53 patients entered into the artificial neural network consisted of a training group of 26 patients, and a testing group of 27 patients. Among these 27 patients, the sensitivity, specificity, and accuracy of the artificial neural network for detecting cardiac death during the follow-up period was 100%, 55%, and 67%, respectively (Table II).

Expert network: To estimate the incremental benefit of the artificial neural network on the risk stratification algorithm, results from the testing group of 27 patients were generalized to the entire 53 patients with a reversible or fixed defect on thallium-201 stress testing. The addition of the neural network to the expert system, creating an expert network, would enable an additional 21 patients to be placed in the low-risk category. The addition of the artificial neural network to the expert system alone increased the accuracy of risk stratification from 78% to 90% ($p < 0.001$). This constituted our expert network (Table II). The overall accuracy of the expert network in the clinical high-risk patients without the 26 test patients (84%) was higher than the expert system alone without the same 26 test patients (68%, $p < 0.05$).

DISCUSSION

Expert systems: One limitation of expert systems is that they have difficulty in modeling chaotic biologic systems. This is because linear mathematic equations don't allow representation of irregular surfaces.[13] The strength of an expert system, on the other hand, is clearly shown by our data. Of the 136 patients classified as low risk for cardiovascular disease by our expert system, only 3 died from cardiac causes over a 4-year period. This degree of accuracy for identification of the low-risk patient is hard to improve upon, whether it be by an artificial neural network, a different expert system, or clinical intuition.

Artificial neural networks: Artificial neural networks have a different set of strengths and weaknesses. A major problem among artificial neural networks is overtraining.[14] When an artificial neural network is overtrained, it models the test group so well that it becomes poor at predicting outcomes when new cases are presented. We chose to train on $\frac{1}{2}$ of the cases, then test only once on the other $\frac{1}{2}$. Testing just once after training seemed a reasonable method to get an accurate gauge of the predictive power of our trained neural network. Use of this method meant that we could only estimate the effect a fully trained neural network would have on all 53 patients. This assumption, however, was not necessary to show the usefulness of the neural network. Without any assumptions, the neural network still demonstrated a statistically significant ability to risk stratify the test group of 27 patients. However, the large number of statistical tests performed in this study could potentially give rise to an incorrect conclusion.

Expert networks: Our study showed that neural network technology can enhance an expert system's ability to risk stratify renal transplant candidates. Further prospective validation of this neural network approach to risk stratification in renal transplant candidates with a larger multicenter study may be useful. At a time when cost control in the medical field is of increased importance, artificial intelligence tools have special appeal because the patient does not need to undergo further testing in order to gain improved risk stratification. The addition of a neural network to an expert system based on clinical variables and thallium-201 stress test data significantly increased the cardiac risk stratification of renal transplant candidates. In this clinical situation, artificial intelligence using both deductive and inductive reasoning—the expert network—appears to be superior to either an expert system or an artificial neural network alone.

Acknowledgment: We thank Kent Spackman, MD, PhD, Director of Health Informatics, Oregon Health Sciences University, for sharing his knowledge of and interest in artificial neural networks.

1. United Network of Organ Sharing. Waiting list by organ by ABO blood group, gender, age, race, and UNOS region. UNOS Update 1996;12:34–35.
2. U.S. Renal Data System, USRDS 1995 Annual Data Report. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD: April 1995.
3. Lawrence J. Introduction to neural networks. 6th ed. Nevada City, CA: California Scientific Software Press, 1994;41–61.
4. Boone JM, Gross GW, Greco-Hunt V. Neural networks in radiologic diagnosis: I. Introduction and illustration. *Invest Radiol* 1990;25:1012–1016.
5. Guerrier MRJ, Detsky AS. Neural networks: what are they? *Ann Intern Med* 1991;115:906–907.
6. 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* 1993;187:81–87.
7. Caudill M. Expert networks. *Byte* 1991°t:108–116.
8. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986;323:533–536.
9. Le A, Wilson R, Douek K, Pulliam L, Tolzman D, Norman D, Barry J, Bennett W. Prospective risk stratification in renal transplant candidates for cardiac death. *Am J Kidney Dis* 1994;24:65–71.
10. Hegewald MG, Palac RT, Angello DA, Perlmutter NS, Wilson RA. Ribose infusion accelerates thallium redistribution with early imaging compared with late 24-hour imaging without ribose. *J Am Col Cardiol* 1991;18:1671–1681.
11. Snedecor GW, Cochran WG. Statistical Methods. 8th ed. Ames, IA: Iowa State University Press, 1989:125–128.
12. Dawson-Saunders B, Trapp RG. Basic and Clinical Biostatistics. 2nd ed. Norwalk, CT: Appleton and Lange, 1994:148–149.
13. McGonigal MD, Cole J, Schwab W, Kauder DR, Rotondo MF, Angood PB. A new approach to probability of survival scoring for trauma quality assurance. *J Trauma* 1993;34:863–870.
14. Doyle HR, Dvorchik I, Mitchell S, Marino IR, Ebert FH, McMichael J, Fung JJ. Predicting outcomes after liver transplantation. A connectionist approach. *Ann Surg* 1994;219:408–415.