British Journal of Psychiatry (1996), **169**, 64–67

Artificial Neural Network to Assist Psychiatric Diagnosis

YIZHUANG ZOU, YUCUN SHEN, LIANG SHU, YUFENG WANG, FENG FENG, KEQIN XU, YING QU, YANMING SONG, YIXIN ZHONG, MINGHUI WANG and WEIQUAN LIU

Background. Artificial Neural Network (ANN), as a potential powerful classifier, was explored to assist psychiatric diagnosis of the Composite International Diagnostic Interview (CIDI).
Method. Both Back-Propagation (BP) and Kohonen networks were developed to fit psychiatric diagnosis and programmed (using 60 cases) to classify neurosis, schizophrenia and normal people. The programmed networks were cross-tested using another 222 cases. All subjects were randomly selected from two mental hospitals in Beijing.
Results. Compared to ICD–10 diagnosis by psychiatrists, the overall kappa of BP network was 0.94 and that of Kohonen was 0.88 (both $P<0.01$). In classifying patients who were difficult to diagnose, the kappa of BP was 0.69 ($P<0.01$). ANN-assisted CIDI was compared with expert system assisted CIDI (kappa=0.72–0.76); ANN was more powerful than a traditional expert system.
Conclusion. ANN might be used to improve psychiatric diagnosis.

The Artificial Neural Network (ANN) is a computing system which mimics the structure of the brain to process information. ANN is composed of many non-linear computational elements operating in parallel and arranged in patterns of biological neural networks, which may carry out many 'intelligent' tasks. ANN is intended to imitate human neural networks in both structure and functioning.

Diagnosis in psychiatry involves non-linear, non-precise, experienced, intuitive, simulative and 'fuzzy' information. While traditional statistics and expert systems are not sensitive, accurate and convenient enough to assess these kinds of data, the ANN, as a new method in classification, prediction and decision-making, is especially powerful in dealing with them. So far, there has been no systematic research applying the ANN to assist psychiatric diagnosis.

To establish a diagnostic tool for worldwide use, the World Health Organization (WHO) and the United States Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA) collaborated in developing the Composite International Diagnostic Interview (CIDI), a full structured diagnostic scale for mental disorders (Robins *et al*, 1988). Using computer programs, the CIDI establishes diagnoses according to both ICD–10 and DSM–III–R (the programs are named CIDI/ICD–10 and CIDI/DSM–III–R; WHO, 1991*c*).

In China, the CIDI was translated into Chinese by the Institute of Mental Health, Beijing Medical University (WHO, 1991*a,b,c*) and field-tested in Chinese by Shu (1993) and Zou *et al* (1994). Zou *et al* (1994) indicated the necessity of improving the diagnostic validity of the CIDI (the kappa fell between 0.33 and 0.79) and listed the possible reasons for the unsatisfactory validity of the CIDI, including non-precise data gathered by the CIDI and a non-flexible, unsuitable computer program.

This study tests the different ANN models and compares the ANN with a traditional expert system in augmenting the diagnosis of the CIDI, to improve the diagnostic validity of the CIDI in relation to clinical diagnoses using ICD–10 and, more generally, establish the possibility and suitability of the ANN in assisting psychiatric diagnosis.

Method

Five psychiatrists in Beijing Huilongguan Hospital were trained to use the CIDI according to the standard procedure of WHO; the kappa of rating consistency among them was 0.902. The CIDI was admin-istered to patients randomly selected from Beijing Huilongguan Mental Hospital and the Teaching Mental Hospital affiliated with Beijing Medical University. The 214 patients included 100 cases of neurosis (obsessive–compulsive disorder, 51; anxiety and phobia, 21; depressive neurosis, 21;

Table 1
General information on participants of study

Group	Age: mean (s.d.)	Education: mean (s.d.)
Normal		
Male (n=47)	34.55 (10.67)	12.06 (3.07)
Female (n=48)	33.58 (8.71)	12.81 (2.51)
Neurosis		
Male (n=62)	30.48 (10.30)	12.48 (2.79)
Female (n=38)	35.00 (13.85)	11.87 (2.96)
Schizophrenia		
Male (n=57)	34.35 (9.83)	11.18 (3.05)
Female (n=57)	32.79 (10.01)	11.68 (3.14)
Summed means		
Male (n=166)	33.12 (10.27)	11.92 (2.97)
Female (n=143)	33.78 (10.85)	12.12 (2.87)

neurasthenia, 6; and hypochondriasis, 1) and 114 cases of schizophrenia. Seventy-four per cent of the patients with neurosis and 97% of the patients with schizophrenia were in-patients; the rest were from out-patient clinics. The course of schizophrenia where applicable was longer than six months (in order to fit the DSM criterion). To establish an outcome diagnosis, all of the patients were diagnosed independently according to ICD–10 by at least two senior psychiatrists who did not know about the research. When the psychiatrists were unanimous on a diagnosis, the patient was included; otherwise the patient was excluded. Ninety-five people in the Beijing area who were not patients were randomly selected as the comparison group. General information about the groups is given in Table 1.

The intelligence levels of all those included in the study were within the normal range (tested by Section M of the CIDI).

Two sets of imitating programs of the ANN, BP network and Kohonen network, were developed to meet the needs of psychiatric diagnosis. All the ANN programs were structured in Borland C++.

In order to orient to the structure of the CIDI and the needs of psychiatric diagnosis, the ANN programs were reformed and modified repeatedly. The structure, training models, initial weights, learning rate, training coefficient, hidden knots and fuzzy weights of the neural networks were set up, tested and adjusted many times. All the procedures reported in this paper have been double-tested to assure stability and optimal functioning of the ANN.

All 396 items from sections B to X of the CIDI were applied as the input of the neural networks. The output of the networks were three categories: normal (control), neurosis, and schizophrenia.

The subjects were divided into two groups. The first group ($n=60$) was used as the training group, with 20 people randomly selected from each of the three categories. The second group, 222 people, was used as the cross-validation group, with 74 drawn from each of the three categories.

The BP network and Kohonen network were separately programmed using the training group, and then both networks were parallel-tested by the cross-validation group.

Results

Diagnostic validity of the BP network

The back-propagation network (BP model) is a 'teacher-indicated' neural network, which means it needs to be 'trained' by 'teachers' represented here by the senior psychiatrists. The network output was compared with the diagnosis established by the senior psychiatrists and the difference was sent back to modify the network by changing the weights of the nodes in the neural network. The network was 'trained' to identify normal, neurosis and schizophrenia by the training group.

Throughout the experiment, the initial weights of the network were set randomly between -0.3 to 0.3, the learning rate was 0.15, the training coefficient (alpha) was 1, and the maximum error level was set as 0.15, 0.1 and 0.05. These parameters optimised the BP network in diagnosis and made the results easy to compare.

The nodes in the hidden layer of the BP network were tried from 20 to 60, and the optimum 40 nodes of the hidden layer were selected for use throughout the experiment.

The model of 'group training', in which the network was 'trained' to learn three categories concurrently rather than one at a time (i.e. 'single case training'), was found to be more efficient and accurate.

The fuzzy weights were used to modify the input variables: 1 (no symptoms) kept 1; 2 (non-clinical meaning) kept 2; 3 (substance misuse) kept 3; 4 (physical problems) kept 4; 5 (mental problems) became 9; 6 (possible), 8 (unclear), and 9 (refused answer) became 2. The inputs fell within 1 to 9, and were ordered to more closely approximate clinical experience.

Table 2 compares the diagnosis of the BP network with the clinical diagnosis made by senior psychiatrists using the cross-validation group.

Diagnostic validity of the Kohonen network

The Kohonen network, unlike the BP, is a self-organised and two-layer neural network, which trains itself without knowing the diagnosis of the training group. To find the optimum learning rate

Table 2
Cross-validity of BP diagnosis (n=222)

Group	BP diagnosis		
	Normal	Neurosis	Schizophrenia
Normal	73	1	0
Neurosis	7	67	0
Schizophrenia	2	0	72

Overall correct rate=95.50%, kappa=0.94, $P<0.01$.

in network training, different learning rates were explored: 0.6, 0.7, 0.8 and 0.9. Throughout the experiment, the optimum learning rate was set as 0.6. The training group and the cross-validation group were the same as in the BP network experiment. Table 3 compares the Kohonen diagnosis with the clinical diagnosis made by senior psychiatrists.

The results show that the kappa for the BP network is higher than for the Kohonen network in assisting the diagnosis of the CIDI: thus the BP network was selected for installation in the CIDI artificial neural network diagnostic system (CIDI–ANN).

In order to further test the diagnostic ability of the CIDI–ANN, 16 patients (10 with neurosis and six with schizophrenia) who were clinically difficult to diagnose were employed to cross-test the system. These patients either had their diagnoses changed during treatment or could not be diagnosed during the first two weeks of their admission to the hospitals. Comparing the diagnosis of the CIDI–ANN with the final diagnosis made by psychiatrists, the overall correct rate was 81.25%, kappa=0.69 and $P<0.01$. Table 4 compares the validity of CIDI/ICD–10, CIDI/DSM–III–R (Zou *et al*, 1994) with the CIDI–ANN in distinguishing neurosis, schizophrenia and normal people. All the results were compared with the standard diagnosis made by psychiatrists according to ICD–10.

The diagnostic validity of the CIDI/ANN is higher than the CIDI/ICD–10 and the CIDI/DSM–III–R. The difference between these systems is more obvious when difficult-to-diagnose patients were tested.

Table 3
Cross-validity of Kohonen diagnosis (n=222)

Group	Kohonen diagnosis		
	Normal	Neurosis	Schizophrenia
Normal	74	0	0
Neurosis	18	56	0
Schizophrenia	3	1	70

Overall correct rate =90.09%, kappa=0.88, $P<0.01$.

Table 4
Comparison of CIDI computer programs

	CIDI/ICD–10	CIDI/DSM–III–R	CIDI–ANN
Neurosis			
Cases	100	100	74
Sensitivity	67%	74%	90.54%
Specificity	99.52%	98.56%	99.34%
Correct rate	89%	90.61%	96.40%
Kappa	0.73	0.77	0.92
Schizophrenia			
Cases	114	114	74
Sensitivity	88.59%	85.96%	97.30%
Specificity	100%	100%	100%
Correct rate	95.79%	94.82%	99.10%
Kappa	0.91	0.89	0.98
Classify neurosis, schizophrenia and controls			
Cases	309	309	222
Correct rate	84.79%	86.41%	95.50%
Overall kappa	0.77	0.79	0.94
Difficult-to-diagnose cases			
Cases	16	16	16
Correct rate	68.75%	50%	81.25%
Kappa	0.48	0.33	0.69

Discussion

Janca *et al* (1992*a*) reported the results of evaluating 20 cases in the United States using CIDI/DSM–III–R. The overall correct kappa was 0.78, with a kappa of 0.84 for depression and 0.76 for anxiety and phobia. Another report by Janca *et al* (1992*b*) used the CIDI/ICD–10 to evaluate 20 American patients and reported an overall correct kappa of 0.77 (0.78 for depression and 0.73 for anxiety and phobia). Comparing the diagnostic validity of the CIDI in China with the United States, the CIDI had similar validity in the two countries.

The CIDI–ANN improved the diagnostic validity of the CIDI (overall correct kappa=0.94). These comparisons indicate that cultural difference did not obviously affect the validity of the CIDI, but the computer method may improve the diagnostic validity of the CIDI.

The CIDI/ICD–10 and CIDI/DSM–III–R were programmed using a logical decision tree which is very limited when dealing with non-linear, non-precise, experienced, intuitive and fuzzy data. Compared with the logical decision tree, the ANN has the following advantages:

(a) The logical decision tree is a one-dimensional-analysis. The ANN is more suitable for analysing and expressing complicated multi-dimensional variables in psychiatric diagnosis.

(b) The ANN may imitate not only logical thought, but also experienced, intuitive and imagery thought. The logical decision tree only imitates logical thought.

(c) The ANN has stronger 'robust' functioning. It is appropriate in managing fuzzy, incomplete and non-precise data found in psychiatric diagnostic scales where many patients cannot give precise and complete answers to all questions of the scales. The logical decision tree often makes diagnostic mistakes when dealing with these kinds of data, because it lacks 'robust' functioning.

Despite the positive results reported in this paper, the CIDI–ANN still has room for improvement. First, it is too long. It is difficult and time consuming for the interviewer to be involved in its administration from beginning to end. It is a particular problem for countries which are short of mental health staff like China. Second, the manually managed data may easily cause errors resulting in low efficiency. Third, more diagnostic categories should be included in the CIDI–ANN.

Although the ANN is a valuable method for assisting psychiatric diagnosis, some of its nature should be noted before it can be widely applied. One is the 'black box' nature of the ANN, which means that the logical procedure of how networks determine a diagnosis cannot be observed. More positively, further development of the ANN can be used to optimise the input variables to compress the length of the CIDI.

Our research indicates that the ANN can improve the diagnostic validity of the CIDI and might become an applicable and helpful method in neuropsychiatry, although further research is necessary.

Acknowledgements

Thanks to Paul Peterson PhD and Sheku G. Kamara PhD of the University of Washington, Washington Institute, for their assistance and opinions.

Clinical implications

- ANN can be used as a powerful classifier in psychiatry.
- ANN procedures can improve the validity of diagnostic scales.
- ANN has higher validity than an expert system in assisting CIDI diagnosis.

Limitations

- Few diagnostic categories used in study.
- ANN was not compared with other statistical procedures such as discriminant and cluster analysis.
- The 'black box' nature of ANN limits understanding of the underlying relationships.

References

Janca, A., Robins, L. N., Bucholz, K. K., *et al* (1992*a*) Comparison of Composite International Diagnostic Interview and Clinical DSM–III–R criteria Checklist Diagnosis. *Acta Psychiatrica Scandinavica*, **85**, 440–443.

——, ——, Cottler, L. B., *et al* (1992*b*) Clinical observation of assessment using the Composite international Diagnostic interview, CIDI Analysis of the CIDI Field Trials – wave II at the St Louis site. *British Journal of Psychiatry*, **160**, 815–888.

Robins, L. N., Wing, J., Wittchen, H. U., *et al* (1988) The Composite International Diagnostic Interview. *Archives of General Psychiatry*, **45**, 1069–1077.

Shu, L. (1993) Field Test of Chinese Version of CIDI–Core Reliability and Validity. *Chinese Mental Health Journal*, **7**, 208–211.

World Health Organization (1991*a*) *CIDI – Core Version 1.0 Interviewer's Copy*. Translated by Institute of Mental Health, Beijing Medical University.

—— (1991*b*) *The CIDI User's Manual*. Translated by Institute of Mental Health, Beijing Medical University.

—— (1991*c*) *The CIDI Computer Manual*. Translated by Institute of Mental Health, Beijing Medical University.

Zou, Y., Shen, Y., Shu, L., *et al* (1994) *Research on Methodology of Psychiatric Diagnosis – Artificial Neural Networks and Diagnostic Scales*. Dissertation of Beijing Medical University.

Yizhuang Zou, MD, **Yucun Shen**, MD, **Liang Shu**, MD, **Yufeng Wang**, MD, Beijing Medical University, Institute of Mental Health, 100083; **Feng Feng**, MD, **Keqin Xu**, MD, **Ying Qu**, MD, **Yanming Song**, MD, Beijing Huilongguan Hospital, 100085; **Yixin Zhong**, PhD, **Minghui Wang**, PhD, **Weiquan Liu**, MS, Beijing Posts and Telecommunication University, China, 100088

Correspondence: Yizhuang Zou, Washington Institute, University of Washington, 9601 Steilacoom Blvd SW, Tacoma, WA 98498-7213, USA. Fax: (206) 756-3987; e-Mail: zou@u.washington.edu

(First received 18 December 1995, final revision 4 March 1996, accepted 21 March 1996)