J Vet Clin 23(2) : 105-110 (2006)

Journal of Veterinary Clinics
www.ksvc.or.kr

Development an Artificial Neural Network to Predict Infectious Bronchitis Virus Infection in Laying Hen Flocks

Son-il Pak[1] and Hyuk-moo Kwon

Department of Veterinary Medicine and Institute of Veterinary Science, Kangwon National University, Chunchon, 200-701, Korea

(Accepted: April 13, 2006)

Abstract : A three-layer, feed-forward artificial neural network (ANN) with sixteen input neurons, three hidden neurons, and one output neuron was developed to identify the presence of infectious bronchitis (IB) infection as early as possible in laying hen flocks. Retrospective data from flocks that enrolled IB surveillance program between May 2003 and November 2005 were used to build the ANN. Data set of 86 flocks was divided randomly into two sets: 77 cases for training set and 9 cases for testing set. Input factors were 16 epidemiological findings including characteristics of the layer house, management practice, flock size, and the output was either presence or absence of IB. ANN was trained using training set with a back-propagation algorithm and test set was used to determine the network's capability to predict outcomes that it has never seen. Diagnostic performance of the trained network was evaluated by constructing receiver operating characteristic (ROC) curve with the area under the curve (AUC), which were also used to determine the best positivity criterion for the model. Several different ANNs with different structures were created. The best-fitted trained network, IBV_D1, was able to predict IB in 73 cases out of 77 (diagnostic accuracy 94.8%) in the training set. Sensitivity and specificity of the trained neural network was 95.5% (42/44, 95% CI, 84.5-99.4) and 93.9% (31/33, 95% CI, 79.8-99.3), respectively. For testing set, AUC of the ROC curve for the IBV_D1 network was 0.948 (SE=0.086, 95% CI 0.592-0.961) in recognizing IB infection status accurately. At a criterion of 0.7149, the diagnostic accuracy was the highest with a 88.9% with the highest sensitivity of 100%. With this value of sensitivity and specificity together with assumed 44% of IB prevalence, IBV_D1 network showed a PPV of 80% and an NPV of 100%. Based on these findings, the authors conclude that neural network can be successfully applied to the development of a screening model for identifying IB infection in laying hen flocks.

Key words **:** chicken infectious bronchitis, neural network, diagnostic accuracy, ROC curve.

Introduction

Avian infectious bronchitis virus (IBV) is known to cause respiratory distress, nephritis, and reproductive disorders including reduced egg production and decreased quality in both layer and broilers (5). It has been reported that clinical signs vary depending on the age of chicks, virulence of the virus strain and existing level of immunity (13,17,25).

Since 1990, the disease continues to be an economically significant problem with a continuing high incidence of IBV infection in domestic laying flocks despite the worldwide use of Mass type infectious bronchitis (IB) vaccines (12,21). Considering a highly contagious characteristics of the disease, the most important aspect for field veterinarians is to making prompt diagnosis of IB to curb further transmission. There are no clear-cut techniques for IBV detection or typing. Since the level of success in detection of IBV after a disease outbreak is influenced by a number of factors (6), detection of IBV from field samples is often unsuccessful. In addition, as genetic variants of IBV have been reported to increasing trend in Korea due to widespread use of live attenuated vaccines (21), classical group-specific ELISA test with reportedly low sensitivities could be considered a very limited practical tool, especially for using this test as a IB control strategy.

From a disease control point of view, prompt prediction of infectious diseases in large-scaled chicken flocks that are likely to have severe disease is very important so that control program can be instituted as early as possible. In this study, this task was investigated using an artificial neural network (ANN). Conventional statistical or modeling approaches using linear models have been employed to determine whether statistically significant correlations exist among independent variables (8). Because field data from flocks with a wide variety of information may be incomplete, it is often difficult to create useful mathematical models for prediction of laboratory or clinical outcomes based on this fuzzy data. ANNs are computer models composed of parallel, non-linear computational elements arranged in layers with a structure that mimics the human brain (2). Their clinical applications include bovine

[1]Corresponding author.
E-mail : paksi@kangwon.ac.kr

mastitis (11), ascites in broilers (18), microbiology (20), laboratory data (1,4,9), radiology (24), and surgery (19). The aim of this study was to develop an ANN model to predict the IB infection with a high degree of accuracy among laying hen flocks with maximum sensitivity by combining both clinical and laboratory data.

Materials and Methods

Study setting and data collection

A comprehensive surveillance program on IB infection (IBP) was carried out since 2003. The purpose of IBP, target population, and farm- and flock-level factors were described earlier (16). Briefly, a total of 86 commercial laying-hen flocks (13 from Kangwon province, 20 Chungcheong province, 25 Gyunggi province and 28 Kyungsang province) were included between May 2003 and November 2005. For practical purposes, flocks in different houses on a single farm and flocks housed in the same house but at different times were considered separate flocks. Data from each study flock were recorded on IBP surveillance form and then entered into a data base.

Serology

The diagnosis of IB was assessed by using clinico-pathological findings and hemagglutination inhibition (HI) test. Detailed procedures were described previously (7,16,23). In this study, flocks with geometric mean HI titer greater than equal to 11 were considered to represent a IB-positive whereas flocks with geometric mean HI titer less than 11 were considered to represent a IB-negative. Because, in Korea, laying hen birds older than 40 weeks of age usually do not vaccinated with killed-virus oil emulsion vaccines, it is very hardly seen this high level of titer among these birds, except in natural infection. Therefore, although the choice of titer 11 as the threshold was a bit arbitrary, but this level of titer was generally acceptable to distinguish presence or absence of infection among flocks.

Construction of the ANN and neurons

A three-layered perceptron ANN was constructed by using a back-propagation software (BrainMaker Professional 3.7 for Windows, California Scientific Software, Nevada City, CA): the input layer propagated to neurons in hidden layer, and then to neuron in output layer, producing a value interpreted as the probability of IB. Specifically, the input layer included 16 neurons, which corresponded to each of the recorded epidemiological characteristics. The schematic architecture of the ANN are shown in Fig 1. For entry into the ANN, quantitative data were normalized to a value over a range between 0 and 1. This was performed by either simply subtracting the lowest value from the original value, then dividing the results by the highest value or simply dividing original value by range of the variable (14). The input neurons, its definition and coding system used for neural network are shown in Table 1. The output of neural network is a continuous variable ranging

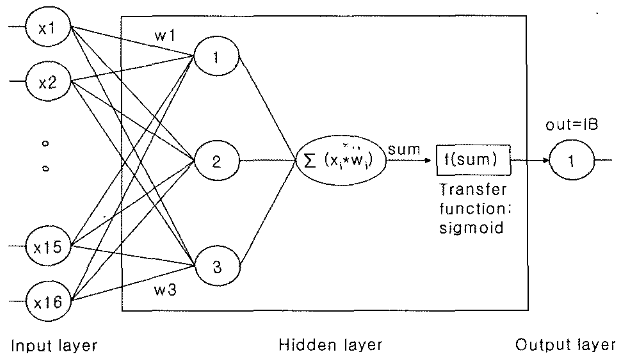

Fig 1. A three-layered fully interconnected neural network architecture. Lines between nodes represent synaptic connections, which are initially determined by random values and then altered by learning rule of the back-propagation algorithm. Sixteen neurons representing presumed risk factors associated with IB form the input layer (left). Each input neuron is synapsed to 3 neurons, forming a single hidden layer (center box). Each hidden layer neuron, in turn, stimulates a single neuron in the output layer (right), producing a normalized value interpreted as the probability of chicken infectious bronchitis virus infection.

from 0 to 1, with 0 being complete certainty by the neural network that IB is absent and 1 representing complete certainty of its presence. In building network, many different networks were trained with different combinations of inputs and training parameters. Initially four different ANNs were considered either by changing the number and characteristics of the input neurons. The first network, IBV_D1, contained 16 input neurons represented by all categorical predictors, the second network, IBV_D2 had 16 normalized predictors including operation in years, flock number and flock size, the remaining two networks, IBV_D3 and IBV_D4 had 13 input neurons with the same types of predictors used for IBV_D1 and IBV_D2, respectively (Table 2).

Training and testing of the ANN

The network was trained by means of the back-propagation rule with a learning rate of 1.0, a smoothing factor of 0.9, and an error training tolerance of 0.1. To help avoid over-training, the mean square errors were monitored at each run to see whether as the run numbers get larger, the error value was decreased. Finally, the point on the error-rate graph at which further training did not decrease the error rate was visually selected. The testing tolerance was set to 0.2. A successful network can generalize and can recognize variations of the cases that it was not exposed to during the training process. In this study, 90% (77 cases) of the total were used to train the network, and the remaining 10% (9 cases) were used in testing (14). Because neural networks cannot learn and fail to converge if data are grouped or ordered in a way. To avoid this problem, before making training sets the data was shuffled once to ensure randomization.

Table 1. Input neurons, its definition and coding system used to train the artificial neural network for predicting chicken infectious bronchitis virus infection

Neuron	Input	Definition	Coding system	
			Category	Values
1	op_year	Farm operation in years	> 5	1.0
			≤5	0.0
2	flock	No. of flocks in a farm	Continuous	Normalized, 0-1
3	size	Size of flocks	< 40,000	0.0
			40,000 - 100,000	0.5
			> 100,000	1.0
4	grw	Type of farm	Growth only	0.0
			Adult only	1.0
5	m_age	Presence of multi-age flocks	Presence	1.0
			Absence	0.0
6	molt	Molting practice	Yes	1.0
			No	0.0
7	introduction	All-in-all-out introduction of flock	Yes	1.0
			No	0.0
8	temp	Checking temperature/humidity	Yes	1.0
			No	0.0
9	grw_type	Type of house	Open	1.0
			Closed	0.0
10	vent	Power ventilation system	Yes	1.0
			No	0.0
11	feces_tr	In-site disposal of feces	Yes	1.0
			No	0.0
12	feces_tn	Frequency of feces disposal	Longer than weekly	1.0
			At least weekly	0.0
13	carcass 1	Incineration of carcass	Yes	1.0
			No	0.0
14	carcass 2	Animal feeds of carcass	Yes	1.0
			No	0.0
15	visitor	Sanitary measure for visitors	Yes	1.0
			No	0.0
16	cold	Presence of cold stress*	Yes	1.0
			No	0.0
17	IB	Presence of IB infection	Continuous	

*This factor was based on the subjective assessment of a veterinarian who is responsible for routine check-up the farm. Yes represents a farm with improper management of in-house temperature control.

Table 2. Four different artificial neural networks and corresponding input neurons

Network	No. input neurons	Characteristics
IBV_D1	16	All categorical
IBV_D2	16	Categorical except for op_year, flock, size
IBV_D3	13	All categorical
IBV_D4	13	Categorical except for op_year, flock, size

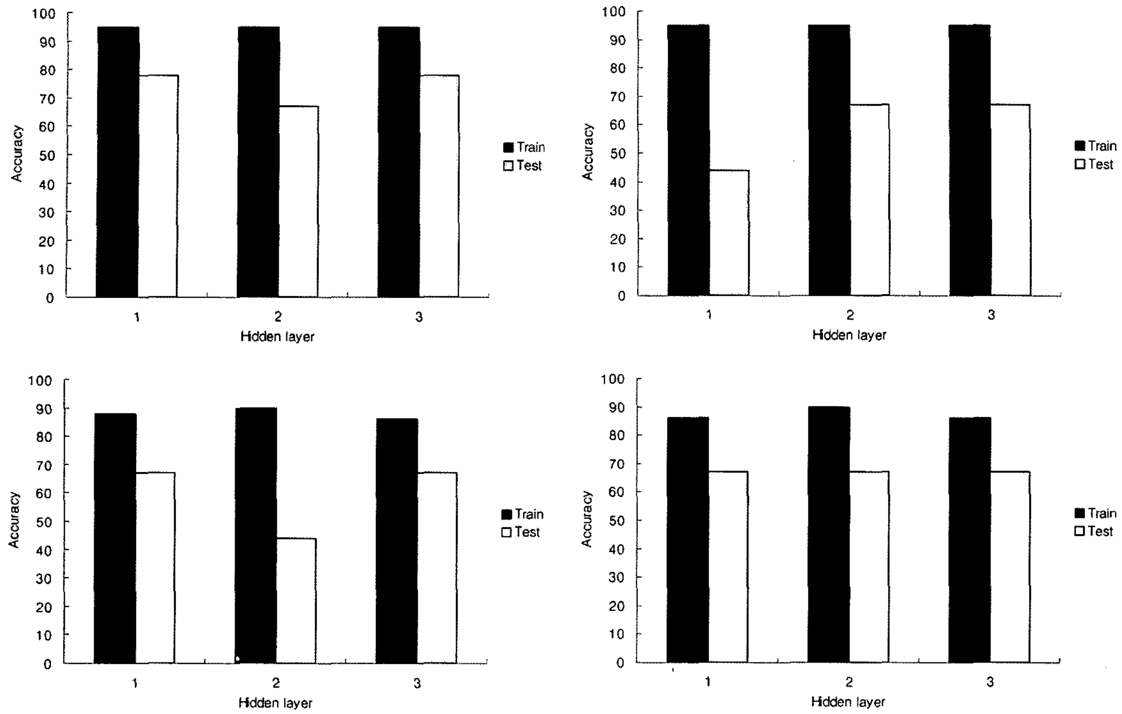

Fig 2. Diagnostic accuracy of the 4 different artificial neural networks (top left, IBV_D1; top right, IBV_D2; bottom left, IBV_D3; bottom right, IBV_D4) for predicting the diagnosis of infectious bronchitis in laying hen flocks. Tolerance parameter was set to 0.1 and 0.2 for training and testing set, respectively.

Evaluation of diagnostic performance of the network

Because the network was trained to a 10% error tolerance, output of the ANN ranged continuously from 0 to 1, with greater output values corresponding to a higher likelihood of being a positive case: output probabilities of less than 0.1 were interpreted as non-IB, probabilities of 0.9 or greater were interpreted as IB, and all intermediate values were interpreted as intermediate by network analysis. However, instead of choosing a particular decision threshold above which a diagnosis of IB would be made, we computed the diagnostic accuracy of network over a range of decision thresholds from 0.05 to 0.95, in increments of 0.05, and at 0.99. We then constructed receiver operating characteristic (ROC) curves by evaluating the number of correctly classified cases obtained at different decision thresholds of the IB output neuron (3,22). The point on the ROC curve corresponding to the best diagnostic accuracy was chosen as the optimal diagnostic sensitivity and specificity. The area under the curve (AUC) was calculated with a software (MedCalc version 7.3, Belgium) (10). The best network for any given combination of training parameters was determined by a model producing the best results on the both training and test set.

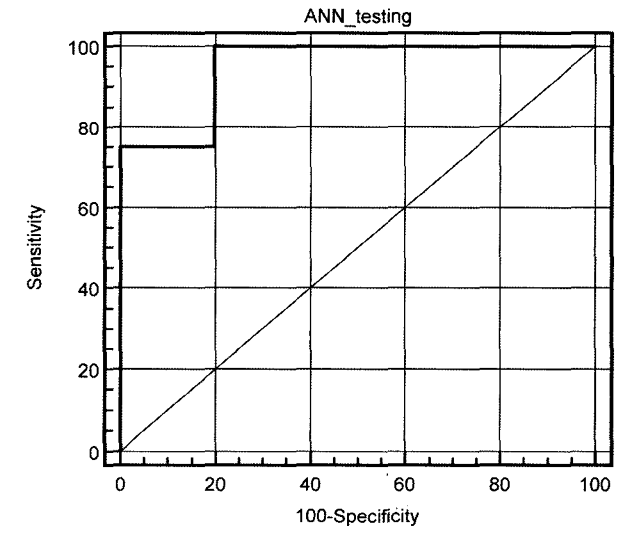

Fig 3. Receiver operating characteristic (ROC) curve for the artificial neural network. The area under the ROC curve was 0.948 (SE=0.086, 95% CI 0.592-0.961).

Results

Diagnostic accuracy of the four different artificial neural networks for predicting the diagnosis of IB in laying hen flocks are shown in Fig 2. Of these, the best network was IBV_D1 with diagnostic accuracy of 94.8% in training tolerance parameter of 0.1. Only the results of the test set of the most accurate network are plotted in ROC curve. ROC curve for performance of the trained IBV_D1 network is depicted in Fig 3. The AUC of the network was 0.948 (SE=0.086, 95% CI 0.592-0.961). At a criterion of 0.7149 or 0.9737, the

Table 3. ANN performance at different criterions for the output value

Criterion(%)	Sensitivity(%)	Specificity(%)	PPV(%)	NPV(%)	Diagnostic accuracy(%)
> 0.1094	100	60	66.7	100	77.8
> 0.7149	100	80	80.0	100	88.9
> 0.9656	75	80	75.0	80.0	77.8
> 0.9737	75	100	100	83.3	88.9
> 0.9981	50	100	100	71.4	77.8

PPV, positive predictive value; NPV, negative predictive value.

diagnostic accuracy was the highest with a 88.9%, but sensitivity was higher in the former criterion (Table 3).

Discussion

Our primary purpose of this study was to develop a network model which can be applied as a screening tool with maximum possible sensitivity in order to identify as many cases of IB event as possible. The IBV_D1 was the best-fitted model and was able to recognize 73 cases out of 77 events, with a sensitivity of 95.5% and a specificity of 93.9% in training set. When applied the model to testing set, diagnostic accuracy (AUC) was 94.8% (overall misclassification of 5.2%). At a criterion of 0.7149 which corresponds to sensitivity of 100%, the IBV_D1 model had a specificity of 80%. The prevalence of IB was 44% (4/9) in testing set and at least to the author's knowledge, this can be assumed as real rate in field settings. In this population, our reported sensitivity and specificity produced a PPV of 80% and an NPV of 100%. At this hypothetical 44% prevalence, a negative test result would thus strongly rule out IB infection, but a positive test would not be particularly helpful due to 20% of false positive. The sensitivity and NPV was substantially decreased when the criterion was raised. In the field setting, it is important to maximize the sensitivity for detecting potential infection to institute control program or diagnostic procedures for the flocks.

Compare to classical methods, the ANN approach provides a couple of potential advantages (1,2,14). First, the ANN is particularly valuable in defining the relationship among data that are not apparent in classical linear statistical models. Second, the ANN offers a computer-aided model that can easily combine the relevant all inputs by encoding the knowledge as numeric weights between interconnected nodes. In addition, the ability of the ANN to assign variable weights to the input parameters to predict IB infection status during the training process means that the most important of the many clinical and laboratory parameters can be identified. Therefore many input data that have little relevance to the outcome can potentially be discarded, thus simplifying the data collection process. However, it is important to emphasize that the results of this study may not be applicable to populations in locations where the epidemiology of IB differs substantially from the area where this study was conducted. Until the model is tested on a different population set, the study can be viewed only as the first attempt in the use of network model in the diagnosis of IB. Since this study was focused on the diagnosis of respiratory-type IB was studied, application of the model to nephropathogenic IB is not currently recommended.

An essential component of ANN model is to enhance the ability of the neural network to generalize to new population samples. This feature can be affected by many factors including the number of neurons in the hidden layer, the type of connections in the network, and the extent to which the network has been trained (1,4). To achieve this purpose, the authors are planning to integrate additional parameters such as previous history of other disease occurrence which shares the clinical findings of IB infection, vaccination, and other epidemiological findings including time interval between chick introduction and farm management factors. In addition, further works are needed to increase the specificity of the model and thus generalize this model in the variety of field settings.

Acknowledgement

This study was supported by a grant No. 203114-03-1-CG000 from the Technology Development Program for Agriculture and Forestry, Ministry of Agriculture and Forestry, Republic of Korea.

References

1. Astion ML, Wilding P. The application of backpropagation neural networks to problems in pathology and laboratory medicine. Arch Pathol Lab Med 1992; 116: 995-1001.
2. Baxt WG. Improving the accuracy of an artificial neural network using multiple differently trained networks. Neural Comp 1992; 4: 772-780.
3. Beck JR, Shultz EK. The use of relative operating characteristic (ROC) curves in test performance evaluation. Arch Pathol Lab Med 1986; 110: 13-20.
4. Cicchetti DV. Neural networks and diagnosis in the clinical laboratory: state of the art. Clin Chem 1992; 38: 9-10.
5. Cunningham CH. Avian infectious bronchitis: characteristics of the virus and antigenic types. Am J Vet Res 1975; 36: 522-523.
6. Dhinakar Raj G, Jones RC. Infectious bronchitis virus: immunopathogenesis of infection in the chicken. Avian Pathol 1997; 26: 677-706.
7. Eidson CS, Page RK, Fletcher OJ, Kleven SH. Vaccination of broiler breeders with teno-synovitis vaccine. Poult Sci 1979; 58: 1490-1497.
8. Feinstein AR. Multiple logistic regression. In: Feinstein AR

(ed). Multivariable analysis. 1st ed. New Haven: Yale University Press, 1996: 297-300.
9. Furlong JW, Dupuy ME, Heinsimer JA. Neural network analysis of serial cardiac enzyme data. Am J Clin Pathol 1991; 96: 134-141.
10. Hanley JA, McNeil BJ. The meaning and use of the area under the receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29-36.
11. Heald CW, Kim T, Sischo WM, Cooper JB, Wolfgang DR. A computerized mastitis decision aid using farm-based records: an artificial neural network approach. J Dairy Sci 2000; 83: 711-720.
12. Kim JH, Song CS, Mo IP, Kim SH, Sung HW, Yoon HS. An outbreak of nephropathogenic infectious bronchitis in commercial pullets. Res Reports Rural Develop Admin 1992; 34: 28-33.
13. Kusters JG, Niesters HGM, Bleumink-Pyuym NMC, Davelaar FG, Horzinek MC, van der Zeijst BAM. Molecular epidemiology of infectious bronchitis virus in the Netherlands. J Gen Virol 1987; 68: 343-352.
14. Lawrence J. Introduction to neural networks: design, theory, and applications. 6th ed, Luedeking S (Ed). Nevada City, CA, California Scientific Software Press, 1994: 154-155.
15. Meistrell ML. Evaluation of neural network performance by receiver operating characteristic (ROC) analysis: examples from the biotechnology domain. Comput Methods Programs Biomed 1990: 32; 73-80.
16. Pak SI, Kwon HM, Yoon HJ, Song CS, Son YH, Mo IP, Song CY. Risk factors for infectious bronchitis virus (IBV) infection in laying flocks in three provinces of Korea: preliminary results. Korean J Vet Res 2005; 45: 405-410.
17. Parsons D, Ellis MM, Cavanagh D, Cook JKA. Characterization of an infectious bronchitis virus isolated from vaccinated broiler breeder flocks. Vet Rec 1992; 131: 408-411.
18. Roush WB, Kirby YK, Cravener TL, Wideman RF. Artificial neural network prediction of ascites in broilers. Poult Sci 1996; 75: 1479-1487.
19. Schwartz MH, Ward RE, MacWilliam C, Verner JJ. Using neural networks to identify patients unlikely to achieve a reduction in bodily pain after total hip replacement surgery. Med Care 1997; 35; 1020-1030.
20. Shang JS, Lin YE, Goetz AM. Diagnosis of MRSA with neural networks and logistic regression approach. Health Care Manag Sci 2000; 3: 287-297.
21. Song CS, Lee YJ, Kim JH, Sung CW, Lee Y, Izumiya T, Miyazawa T, Jang HK, Mikami T. Epidemiological classification of infectious bronchitis virus isolated in Korea between 1986 and 1997. Avian Pathol 1998; 27: 409-416.
22. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988; 240: 1285-1293.
23. Thayer SG, Eidson CS, Page RK, Kleven SH. Multivalent inactivated virus oil emulsion vaccines in broiler breeder chickens. I. Newcastle disease virus and infectious bursal disease virus bivalent vaccines. Poult Sci 1983; 62: 1978-1983.
24. Tourassi GD, Markey MK, Lo JY, Floyd CE Jr. A neural network approach to breast cancer diagnosis as a constraint satisfaction problem. Med Phys 2001; 28: 804-811.
25. Yu L, Jiang Y, Low S, Wang Z, Nam SJ, Liu W, Kwangac J. Characterization of three infectious bronchitis virus isolates from China associated with proventriculus in vaccinated chickens. Avian Dis 2001; 45: 416-424.

산란계의 전염성 기관지염을 예측하기 위한 인공신경망 모형의 개발

박선일[1] · 권혁무

강원대학교 수의학과 동물의학종합연구소

요 약: 2003년 5월부터 2005년 11월까지 산란계의 전염성기관지염(IB) 예찰 프로그램에 등록한 농장에 대한 역학조사에서 얻은 자료에 근거하여 IB 감염을 확인할 수 있는 모형을 구축하기 위하여 16개의 입력 뉴런, 3개의 은닉 뉴런, 1개의 출력 뉴런으로 구성된 3층 인공신경망 모형을 개발하였다. 총 86개의 계군 중 77개는 훈련자료에 할당하고 나머지 9개는 검정자료로 무작위로 할당하여 back-propagation algorithm으로 신경망 훈련을 수행하였다. 입력 뉴런은 산란계군의 특성, 사양관리, 계군의 크기 등 16개의 역학조사 항목을 사용하였으며 출력 뉴런은 IB 감염의 유무로 투입하였다. 훈련된 신경망을 검정자료에 적용하여 민감도와 특이도를 산출하였으며 진단의 정확도는 receiver operating characteristic (ROC) 곡선을 사용하여 곡선 밑의 면적(AUC)을 계산하여 평가하였다. 입력 뉴런의 특성과 훈련모수를 변경하면서 다양한 신경망을 구성하였으며 최적의 신경망으로 확인된 IBV_D1 신경망의 경우 훈련자료에 대하여 77건 중 73건을 올바르게 판단하여 94.8%의 정확도를 보였다. 민감도와 특이도는 각각 95.5% (42/44, 95% CI, 84.5-99.4)와 93.9% (31/33, 95% CI, 79.8-99.3)로 나타났다. 훈련된 신경망을 검정자료에 적용하여 ROC 곡선을 작성한 결과 AUC는 전체의 94.8% (SE=0.086, 95% CI 0.592-0.961)를 차지하는 우수한 모형으로 나타났다. ROC 곡선에서 기준을 0.7149 이상으로 판단할 때 진단의 정확도가 88.9%로 가장 높았으며 100%의 민감도를 달성하였다. 이러한 민감도와 특이도에서 44%의 IB 유병률을 가정할 때 IBV_D1 모형은 80%의 양성예측도와 100%의 음성예측도를 보였다. 이러한 소견에 근거할 때 본 연구에서 구축한 신경망 모형은 산란계군에서 IB의 존재를 확인하기 위한 목적에 성공적으로 응용될 수 있을 것으로 판단되었다.

주요어: 닭 전염성기관지염, 인공신경망, 진단 정확도, ROC 곡선