ORIGINAL ARTICLE

Artificial neural network analysis for evaluating cancer risk in multinodular goiter

Baris Saylam, Mehmet Keskek, Sönmez Ocak, Ali Osman Akten, Mesut Tez
Department of Surgery, Ankara Numune Training and Research Hospital, Ankara, Turkey

Background: The aim of this study was to create a diagnostic model using the artificial neural networks (ANNs) to predict malignancy in multinodular goiter patients with an indeterminate cytology. **Materials and Methods:** Out of 623 patients, 411 evaluated for multinodular goiter between July 2004 and March 2010 had a fine-needle aspiration biopsy. All patients underwent total thyroidectomy. The interpretation was consistent with an indeterminate lesion in 116 (18.6%) patients. Patient's medical records including age, sex, dominant nodule size, pre-operative serum thyroid-stimulating hormone level, thyroid hormone therapy and final pathologic diagnosis were collected retrospectively. **Results:** The mean age of the patients was 44.6 years (range, 17-78 years). About 104 (89.7%) were female and 12 (10.3%) were male patients. Final pathology revealed 24 malignant diseases (20.7%) and 92 (79.3%) benign diseases. After the completion of training, the ANN model was able to predict diagnosis of malignancy with a high degree of accuracy. The area under the curve of ANNs was 0.824. **Conclusion:** The ANNs technique is a useful aid in diagnosing malignancy and may help reduce unnecessary thyroidectomies in multinodular goiter patients with an indeterminate cytology. Further studies are needed to construct the optimal diagnostic model and to apply it in the clinical practice.

Key words: Artificial neural network, fine-needle aspiration, ındeterminate cytology, multinodular goiter, thyroid

How to cite this article: Saylam B, Keskek M, Ocak S, Akten AO, Tez M. Artificial neural network analysis for evaluating cancer risk in multinodular goiter. J Res Med Sci 2013;18:553-6.

INTRODUCTION

Thyroid nodules constitute the main indication for fine-needle aspiration and the goal of this diagnostic procedure is to detect the thyroid neoplasms that require surgical resection. The diagnosis of a thyroid fine-needle aspiration biopsy (FNAB) is generally divided into three categories: Benign, malignant and indeterminate. The latter group usually includes follicular neoplasms, follicular lesions and sometimes a more specific diagnosis such as Hurthle cell neoplasm or follicular lesion/neoplasm with Hurthle cell change.[1]

Given the frequency of thyroid lesions with an indeterminate diagnosis, it is of considerable clinical importance to refine the pre-operative evaluation of cancer risk in these lesions, in order to identify a low-risk subgroup of patients that could spare thyroidectomy as suggested by previous studies.[2]

In the last 10 years, a class of techniques inspired by the workings of biologic neurons, artificial neural networks (ANNs), has been proposed as a supplement or alternative to standard statistical techniques for predicting complex biologic phenomena. Briefly, ANNs are a class of nonlinear mathematical models that are characterized by a complex structure of interconnected computational elements, the neurons. These computational elements aggregate a series of inputs by using a summation operation and produce an output, such as the prediction of malignancy.[3]

Therefore, the purpose of this study was to develop an ANNs model to prediction of malignancy in multinodular goiter patients with an indeterminate cytology and evaluate the predictive performance of the ANNs model in comparison with the traditional logistic regression (LR) model.

MATERIALS AND METHODS

Study population

Out of 623 patients, 411 evaluated for multinodular goiter between July 2004 and March 2010, had a FNAB. All patients underwent total thyroidectomy. The interpretation was consistent with an indeterminate lesion in 116 (18.6%) patients. These 116 patients comprised the study population. The technique of FNAB and the cytologic criteria followed for smear evaluation have been previously described.[4]

Data collection

All patients underwent total thyroidectomy. Patient's medical records including age, sex, dominant nodule

Address for correspondence: Dr. Barış Saylam, 1443 Cadde 35/10 Kizilirmak Mah., 06510 Çankaya, Ankara, Turkey. E-mail: mdbsaylam@yahoo.com
Received: 14-07-2012; **Revised:** 08-09-2012; **Accepted:** 14-01-2013

size, pre-operative serum thyroid-stimulating hormone level, thyroid hormone level, thyroid hormone therapy and final pathologic diagnosis were collected retrospectively.

Training and validation data sets

Nearly 75% of the patients ($N1$ = 87) were randomly selected to provide the training group for constructing ANN and LR prediction models. The remaining 25% of participants ($N2$ = 29) were assigned to the validation group for performance comparisons of ANN and LR models. The proportion of malignancy between the training group and validation group was similar and there was no statistically significant difference by χ^2 test for gender and age between the training and validation datasets.

Modeling tools

LR

The LR model generates the coefficients for the following formula used in logit transformation for predicting the probability of an event among patients with certain characteristics of interest. Logit (P) = $\beta_0 + \beta_1\chi_1 + \beta_2\chi_2 + \ldots + \beta_i\chi_i$.[5] The formula $P = 1/(1 + e^{-\text{logit (P)}})$ used for calculating the probability of malignancy in this study, where 1 = malignant and 0 = benign. A stepwise algorithm was used to construct the multivariate LR model.

ANN

The ANN is nonlinear statistical data modeling tool that can be used to model complex relationships between inputs and outputs or to find patterns in data. The processing elements or nodes are arranged in "input," "hidden and output" layers, each layer containing one or more nodes. The input layer consists of the data thought to be of value in predicting the outputs of the model. Each data point is represented by a node in the input layer. The output layer estimates the probability of the outcome as determined by the model. Each layer comprises one or more processing elements all interconnected in a way that each node in the hidden layer is connected to each node in the input and output layers. Each connection carries a "weight" or value that determines the relevance of a particular input for the resulting output. The ANNs makes predictions based on the strength of connections between the neurons in the input, hidden and output layers. The results of the output layer in ANNs model represent the probability of a characteristic of interest (malignancy).[5,6] The following information was used for ANNs analysis: Patient age (years), gender (male; female), thyroid function status (euthyroid, hypothyroid or hyperthyroid), previous hormone therapy and ultrasound-evaluated maximum nodule diameter (mm). Three-layered, multilayer perceptron ANNs, with back propagation circuit, were constructed using the IBM SPSS Statistics 19 (IBM Corp., Armonk, NY, USA) Applying back-propagation allows a model that starts with known inputs and random outputs to be trained until the ANNs output values match the expected output. Four "hidden" neurons were used. The ANNs was iterated in excess of 100,000 epochs for the training set of 87 patients, with 89.7% accuracy. Batch learning process was used.

Statistical analysis

Model discrimination was measured by the area under the curve (AUC) the receiver-operator characteristic (ROC) to evaluate how well the model distinguished patients experienced the events from those who did not. An AUC of 0.5 indicates that the model does not predict better than chance. The discrimination of a diagnostic model is considered perfect if AUC is equal to 1, good if AUC is greater than 0.8, moderate if AUC is 0.6-0.8 and poor if AUC is lesser than 0.6.[7]

RESULTS

The mean age of the patients was 44.6 years (range, 17-78 years). About 104 (89.7%) were female and 12 (10.3%) were male patients. Final pathology revealed 24 malignant diseases (20.7%) and 92 (79.3%) benign diseases. Mean age of cancer patients was 51 (range, 19-73). Twenty-one patients were female (87.5%) and three patients were male (12.5%). Of the malignant patients, 14 patients (13.3%) had papillary carcinoma, eight patients (7.6%) had papillary microcarcinoma (papillary carcinoma < 1 cm) and two patient (1.9%) had follicular microcarcinoma.

Comparison of models

The ANNs and LR models could successfully detect an individual's risk of having thyroid cancer. Figures 1 and 2 summarize the ROC curve obtained from the ANNs and LR models. For the ANNs model, the sensitivity, specificity, positive predictive value and negative predictive value were 72%, 94%, 99% and 95%, respectively while the corresponding numbers were 66%, 78%, 70% and 44% in the LR model, respectively [Table 1].

The AUC value of the ANNs model was higher (AUC = 0.824) than that of the LR model (AUC = 0.702) [Table 1].

DISCUSSION

The incidence of the pre-operative cytologic diagnosis of cellular microfollicular among all aspirates in different

Table 1: Performance of ANN and LR models on diagnosis of malignancy

	AUC	SN	SP	PPV	NPV
ANN	0.824	0.72	0.94	0.99	0.95
LR	0.702	0.66	0.78	0.7	0.44

ANN=Artificial neural network; LR=Logistic regression; AUC=Area under ROC curve; SN=Sensitivity; SP=Specificity; PPV=Positive predictive value; NPV=Negative predictive value; ROC=Receiver-operator characteristic

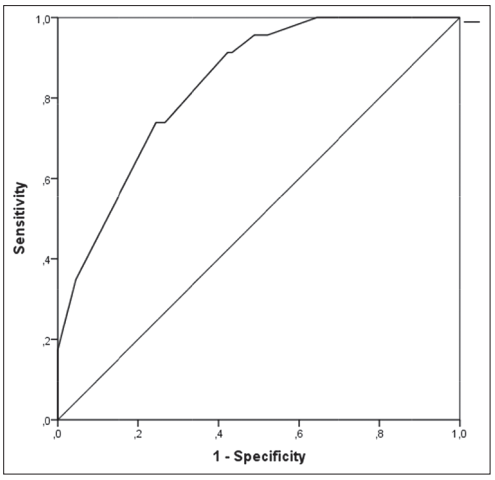

Figure 1: Receiver-operator characteristic analysis of artificial neural networks model

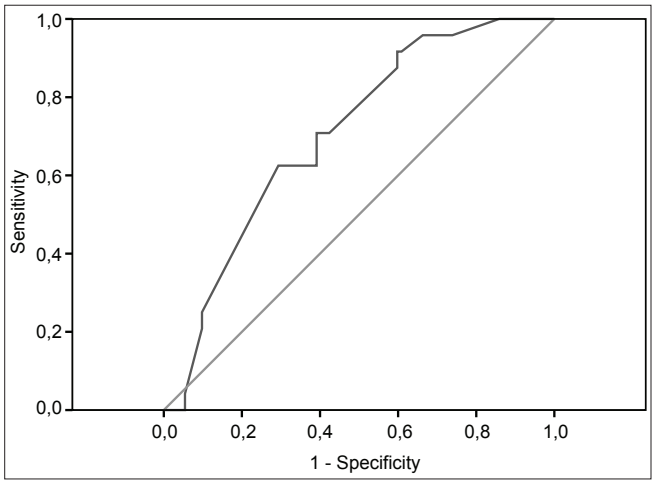

Figure 2: Receiver-operator characteristic analysis of logistic regression model

series varies from 6% to 26%. These include follicular or suspicious lesions, categories in which cancer rate ranges from 10% to 40%. Most clinicians recommend surgery in all these patients because even a cancer risk of 10% is considered as unacceptably high. On the other hand, approximately 80% of the patients underwent unnecessary thyroidectomy.[1,2]

A variety of approaches have been suggested to improve diagnostic accuracy in patients with cytologically indeterminate thyroid lesions. These include cytometric deoxyribonucleic acid measurements, morphometric studies and the use of immunohistochemical markers on cell aspirates. However, none of these approaches has a relevant impact in clinical practice.[2]

ANNs has been proposed as a supplement or alternative to standard statistical techniques for predicting complex biologic phenomena. Briefly, ANNs are a class of nonlinear mathematical models that are characterized by a complex structure of interconnected computational elements, the neurons. These computational elements aggregate a series of inputs by using a summation operation and produce an output, such as the presence of malignancy. The ANNs can be used to perform nonlinear statistical modeling and provide a new alternative to LR, the most commonly used method for developing predictive models for dichotomous outcomes in medicine. The ANNs offers a number of advantages, including requiring less formal statistical training, ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables and the availability of multiple training algorithms.[8,9]

Previously, Ippolito *et al*. studied 453 patients with a thyroid nodule (64 multinodular goiter, 389 solitary nodule) diagnosed as indeterminate at FNAB by using a feed-forward ANNs analysis to integrate cytologic and clinical data, neural network analysis subdivided the 371 lesions of the first series into a high-risk group (cancer rate of approximately 33% at histology) and a low-risk group (cancer rate of 3%). Only cytologic parameters contributed to this classification. Analysis of the ROC curves (AUC: 0.87) demonstrated that the ANNs model discriminated with higher sensitivity and specificity between benign and malignant nodules compared with standard cytologic criteria.[2]

One limiting factor in our study was the restricted data that was derived from a single institution and the small sample size. Our results must be examined on external data obtained at other institutions.

CONCLUSIONS

The ANNs technique is a useful aid in diagnosing malignancy and may help reduce unnecessary thyroidectomies in multinodular goiter patients with an indeterminate cytology. Further studies are needed to construct the optimal diagnostic model and to apply it in the clinical practice.

REFERENCES

1. Akgul O, Ocak S, Gocmen E, Koc M, Tez M. Clinical significance of cellular microfollicular lesions in goiter. Endocrinologist 2010;20:115-6.
2. Ippolito AM, De Laurentiis M, La Rosa GL, Eleuteri A, Tagliaferri R, De Placido S, *et al*. Neural network analysis for evaluating cancer risk in thyroid nodules with an indeterminate diagnosis at aspiration cytology: Identification of a low-risk subgroup. Thyroid 2004;14:1065-71.
3. Sargent DJ. Comparison of artificial neural networks with other statistical approaches: Results from medical data sets. Cancer 2001;91:1636-42.
4. La Rosa GL, Belfiore A, Giuffrida D, Sicurella C, Ippolito O, Russo G, *et al*. Evaluation of the fine needle aspiration biopsy

in the preoperative selection of cold thyroid nodules. Cancer 1991;67:2137-41.
5. Wang CJ, Li YQ, Wang L, Li LL, Guo YR, Zhang LY, *et al*. Development and evaluation of a simple and effective prediction approach for identifying those at high risk of dyslipidemia in rural adult residents. PLoS One 2012;7:e43834.
6. Ho WH, Lee KT, Chen HY, Ho TW, Chiu HC. Disease-free survival after hepatic resection in hepatocellular carcinoma patients: A prediction approach using artificial neural network. PLoS One 2012;7:e29179.
7. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
8. Göçmen E, Klc YA, Yoldaş O, Ertan T, Karaköse N, Koç M, *et al*. Comparison and validation of scoring systems in a cohort of patients treated for biliary acute pancreatitis. Pancreas 2007;34:66-9.
9. Biglarian A, Bakhshi E, Gohari MR, Khodabakhshi R. Artificial neural network for prediction of distant metastasis in colorectal cancer. Asian Pac J Cancer Prev 2012;13:927-30.

Source of Support: Nil, **Conflict of Interest:** None declared.