The Evolution of a Malignancy Risk Prediction Model for Thyroid Nodules Using the Artificial Neural Network

Article *in* Middle East Journal of Cancer · January 2016

CITATIONS
8

READS
172

12 authors, including:

Shahram Paydar
Shiraz University of Medical Sciences
203 PUBLICATIONS **1,896** CITATIONS
SEE PROFILE

Saeedeh Pourahmad
Shiraz University of Medical Sciences
67 PUBLICATIONS **876** CITATIONS
SEE PROFILE

Mohsen Azad
Hormozgan University of Medical Sciences
21 PUBLICATIONS **196** CITATIONS
SEE PROFILE

Shahram Bolandparvaz
Shiraz University of Medical Sciences
109 PUBLICATIONS **1,195** CITATIONS
SEE PROFILE

Original Article
Middle East Journal of Cancer; January 2016; 7(1): 47-52

The Evolution of a Malignancy Risk Prediction Model for Thyroid Nodules Using the Artificial Neural Network

Shahram Paydar*, Saeedeh Pourahmad, Mohsen Azad**, Shahram Bolandparvaz*, Reza Taheri***, Zahra Ghahramani*, Ali Zamani****, Marjan Jeddi****, Fariba Karimi****, Mohammad Hossein Dabbaghmanesh****, Mesbah Shams****, Hamid Reza Abbasi*♦**

*Trauma Research Center, Shahid Rajaee (Emtiaz) Trauma Hospital, Shiraz University of Medical Sciences, Shiraz, Iran
**Department of Biostatistics, Shiraz University of Medical Sciences, Shiraz, Iran
***Department of Neurosurgery, Shiraz University of Medical Sciences, Shiraz, Iran
****Endocrinology and Metabolism Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

Abstract

Background: Clinically frank thyroid nodules are common and believed to be present in 4% to 10% of the adult population in the United States. In the current literature, fine needle aspiration biopsies are considered to be the milestone of a model which helps the physician decide whether a certain thyroid nodule needs a surgical approach or not. A considerable fact is that sensitivity and specificity of the fine needle aspiration varies significantly as it remains highly dependent on the operator as well as the cytologist's skills. Practically, in the above group of patients, thyroid lobectomy/isthmusectomy becomes mandatory for attaining a definitive diagnosis where the majority (70%-80%) have a benign surgical pathology. The scattered nature of clinically gathered data and analysis of their relevant variables need a compliant statistical method. The artificial neural network is a branch of artificial intelligence. We have hypothesized that conduction of an artificial neural network applied to certain clinical attributes could develop a malignancy risk assessment tool to help physicians interpret the fine needle aspiration biopsy results of thyroid nodules in a context composed of patient's clinical variables, known as malignancy related risk factors.

Methods: We designed and trained an artificial neural network on a prospectively formed cohort gathered over a four year period (2007-2011). The study population comprised 345 subjects who underwent thyroid resection at Nemazee and Rajaee hospitals, tertiary care centers of Shiraz University of Medical Sciences, and Rajaee Hospital as a level I trauma center in Shiraz, Iran after having undergone thyroid fine needle aspiration. Histopathological results of the fine needle aspirations and surgical specimens were analyzed and compared by experienced, board-certified pathologists who lacked knowledge of the fine needle aspiration results for thyroid malignancy.

Results: We compared the preoperative fine needle aspiration and surgical histopathology results. The results matched in 63.5% of subjects. On the other hand, fine needle aspiration biopsy results falsely predicted malignant thyroid nodules in 16% of cases (false-negative). In 20.5% of subjects, fine needle aspiration was falsely positive for thyroid malignancy. The Resilient back Propagation (RP) training algorithm lead to acceptable accuracy in prediction for the designed artificial neural network (64.66%) by the cross- validation method. Under the cross-validation method, a back propagation algorithm that used the resilient back propagation protocol - the accuracy in prediction for the trained artificial neural network was 64.66%.

Conclusion: An extensive bio-statistically validated artificial neural network of certain clinical, paraclinical and individual given inputs (predictors) has the capability to stratify the malignancy risk of a thyroid nodule in order to individualize patient care. This risk assessment model (tool) can virtually minimize unnecessary diagnostic thyroid surgeries as well as FNA misleading.

Keywords: Malignancy, Risk prediction model, Thyroid nodules, Artificial neural network

♦Corresponding Author:
Hamid Reza Abbasi, MD
Trauma Research Center, Shahid Rajaee (Emtiaz) Hospital, Shiraz University of Medical Sciences, Shiraz, Iran.
Tel: +98-71-36360697
Fax: +98-71-36254206
Email: abbasimezy@yahoo.com

Middle East Journal of Cancer

Received: April 25, 2015; Accepted: August 12, 2015

Introduction

The lifetime risk for developing a palpable thyroid nodule is estimated to be 5%-10%.[1] Clinically, frank thyroid nodules are quite common and presumed present in 4% to 10% of the adult population in the United States. A number of autopsy surveys have concluded that these figures may probably underestimate the frequency of thyroid nodules.[1] As thyroid nodules reach a more prevalent percentile, physicians are urged to find a more meaningful determinant that can guide them to a more reliable algorithm to assess the risk of thyroid malignancy.[2]

In the current literature, fine needle aspiration (FNA) biopsies are set to be the milestone of a model which helps the physician to decide whether a certain thyroid nodule needs a surgical approach or not.[3,4] Although the FNA biopsy is considered to be the most cost-effective, it remains an imperfect diagnostic tool for thyroid nodules evaluated in current practice.[5]

Current literature classifies the FNA results based on an increasing risk of malignancy across a spectrum composed of six well defined categories:[5]

1) Unsatisfactory or no diagnostic FNA results with an unknown risk of malignancy.

2) Benign results with less than 1% malignancy risk.

3) Follicular lesions of undetermined significance with a 5%-10% malignancy risk.

4) Follicular neoplasm which has a 20%-30% malignancy risk.

5) Results equivocal or suspicious for malignancy (50%-75% malignancy risk).

6) Malignant results (100% malignancy risk).

The identified risk factors for the presence of carcinoma within thyroid nodules include age, gender, prior history of neck irradiation and larger thyroid nodule size.[1] Sensitivity and specificity of the FNA varies significantly as it remains highly dependent on the operator and the cytologist's skills.[5] Tee et al. have reported a sensitivity as low as 66% but this figure approximated 80% in a comprehensive review of the published literature. The researchers claimed that FNA was capable of missing up to a third of all thyroid malignancies.[6]

Stojadinovic et al. reported up to 61% inaccuracy in initial diagnosis when cytology (FNA) was compared to histology or revised later by an expert cytologist.[5] The rate of false-positive and false-negative results attributed to FNA have remained a challenge.[3] However, patients with negative FNA findings were often excluded in statistical analysis and this might lead to bias in the reported sensitivity of FNA.[6]

Over 20% of patients who undergo FNA of a thyroid nodule have an indeterminate cytology with an associated malignancy risk prevalence of 20%-30% as noted earlier. Currently, the thyroid lobectomy/isthmusectomy procedure is mandatory for the above groups of patients in order to attain a definitive diagnosis. However, the majority (70%-80%) have a benign surgical pathology.[5] In addition, FNA biopsies of large thyroid nodules (at least 4 cm) yield a high false-negative rate, therefore the patient becomes a candidate for diagnostic lobectomy.[2] On the other hand, a problem may arise when FNA misses up to a third of all thyroid malignancies.[6]

In order to minimize the morbidity associated with unnecessary diagnostic surgery as well as neglected suspicious cases, there is a mandatory need to establish and evolve a clinical malignancy risk prediction model to assist the clinician with interpretation of an FNA biopsy of thyroid nodules.

Development of a trained system provides users with the ability to train these networks in order to act as smart, logic circuits to estimate the probability of disease or disorder based on input variables that use an intuitive computer program.[5] Stojadinovic et al. have developed an integrated predictive decision model that used a Bayesian Belief Network analytical tool which compiled the joint probability distribution of all variables in the data set by building a directed acyclic network of conditional probabilities. This tool was significantly effective in predicting malignancy in thyroid nodules.[5]

The scattered nature of clinically gathered data and analysis of their relevant variables needs a

Table 1. True positive and negative values for fine needle aspiration (FNA) biopsies versus surgical specimen analyses.

FNA Histopathology	Benign	Malignant	Total
Benign (%)	118 (62.4)	71 (37.6)	189 (100)
Malignant (%)	55 (35.3)	101 (64.7)	156 (100)

compliant statistical method. The artificial neural network (ANN) is a branch of artificial intelligence. It is a nonlinear modeling method that has the capability to find complex relationships and patterns between large amounts of data without being disturbed by multiple variables as inputs. The ANN is derived from brain neuronal architecture which can simulate how the brain behaves when it predicts an antecedent event based on previous input of detailed data.

We hypothesized that application of an ANN to certain clinical attributes would enable development of a malignancy risk assessment tool that could assist physicians with interpretation of FNA biopsy results of thyroid nodules in a context composed of patient`s clinical variables, known as malignancy related risk factors. We further hypothesized that an ANN derived predictive model could potentially produce a feasible decision making tool in order to help physicians predict malignancy in thyroid nodules.

Materials and Methods

This prospective observational cohort study evaluated the diagnostic accuracy of a preoperative malignancy risk assessment tool in patients scheduled to undergo thyroidectomies. We designed and trained an ANN on a prospectively formed cohort gathered over a four year period (2007-2011).

The study population comprised 345 subjects who underwent thyroid resection in Nemazee and Rajaee Hospitals, tertiary care centers of Shiraz University of Medical Sciences, and Rajaee Hospital, a level I trauma center in Shiraz, Iran. Patients previously underwent thyroid FNA biopsies. Experienced, board-certified pathologists who were blinded to the FNA results analyzed and compared the histopathological FNA results to surgical histopathology results. All data were collected via a review of patients' files. This study received local Institutional Review Board approval.

Developed neural network

We introduced the study data to an ANN to analyze and further simulate a modeling method to render the relationship between the inputs. The purpose of this modeling method was to design and train a network which could assess the malignancy risk probability in a patient when certain inputs known as risk factors for thyroid malignancy were given. This would help the physician to interpret the results of FNA biopsies in the context of other risk factors in order to minimize unnecessary diagnostic surgeries and neglect of suspicious cases.

We began the modeling process by randomly dividing subjects into two groups. The larger group consisted of 70% to 75% of patients to train the network (training set); the remainder (testing set) were used to evaluate the trained network with the intent to identify its predictive value. In this study we used back propagation ANN composed of ten neurons in the hidden layer and two neurons in the output layer that were trained via the resilient back propagation (RP) protocol. We further evaluated this neural network by two methods:[7] 1) the receiver-operating characteristic curve (ROC) and 2) cross-validation. Analysis was subsequently performed with MATLAB 7.11.0. The ROC curve method drew the probability of correct diagnosis of malignant nature of a certain thyroid nodule by the network (sensitivity) versus the false diagnosis of a nodule as benign (false-negative value or specificity). A higher area under the curve (AUC) has a better predictive ability. Cross-validation is a technique in which the data set is arbitrarily divided into a number of parts (usually 3 to 10).

All parts, with the exception of one, are used for training; the remaining part is used for testing the network. Next, the accuracy rate of the trained network is calculated. This process is repeated interactively by substituting the testing set by other parts of the data set until all parts are used to test the network. The final accuracy rate is calculated as the average of the accuracy rates from each step.

Results

A comparison of the preoperative FNA and surgical histopathology results indicated that the results matched in 63.5% of subjects. On the other hand, the FNA biopsy falsely predicted a malignant nature for thyroid nodules in 16% of subjects (false-negative); in 20.5% of subjects the FNA was falsely positive for thyroid malignancy (Table 1).

Table 2 shows the study population characteristics that formed the inputs for the ANN. These descriptive data sets included thyroid malignancy risk factors as the current literature mentions. The output of this network was nodule status, benign or malignant. Status was definitely proven by a prospective review of the surgical histopathology results.

According to the cross-validation method, the back propagation artificial neural network with the RP training method had an accuracy of 64.5% for correct prediction. The AUC for the trained ANN was 0.72 (Table 3).

Table 2. Descriptive characteristics of the study population.

Variable	Frequency
Gender	
Male	66 (19.1%)
Female	279 (80.9%)
Presence of multiple nodules	
Yes	182 (52.8%)
No	163 (47.2%)
Rapid growth	
Yes	251 (72.8%)
No	94 (27.2%)
Family history of thyroid disease	
Positive	60 (17.4%)
Negative	285 (82.6%)
Family history of malignancy	
Positive	61 (17.7%)
Negative	284 (82.3%)
Age (years)	
≤36	136 (39.4%)
>36	209 (60.6%)
Size of thyroid gland (cm)	
≤5.5	153 (44.3%)
>5.5	192 (55.7%)
Size of thyroid nodule (cm)	
≤2	230 (66.7%)
>2	115 (33.3%)
Disease duration (years)	
≤1.5	175 (50.7%)
>1.5	170 (49.3%)
Nature of disease	
Benign	189 (54.8%)
Malignant	158 (45.2%)

Discussion

The present study introduced a bioinformatics-derived preoperative malignancy risk assessment model. We conducted a prospective observational cohort study that used this model to evolve an ANN based risk stratification tool which was unique for each patient scheduled for thyroid resection.

Thyroid nodules are quite common findings among patients who seek medical help in the US. However, the majority of palpable nodules are found to be pathologically benign in nature according to surgical histopathology findings.[8] Fine needle aspiration is currently the initial diagnostic workup with a significantly diverse sensitivity and specificity that is highly dependent on the expertise of the operator as well as the cytologist. Of note, more than 20% of indeterminate FNA biopsy results push the surgeon to consider thyroidectomy purely to achieve a clear diagnosis and rule out a questionable malignant nodule.[9]

We have applied the ANN on prospectively gathered clinical data that included paraclinical, clinical, and individual, as well as histopathological predictors for malignancy. Next, we designed, evolved and cross-validated a subject-individualized malignancy risk stratification neural network in order to clarify a patient-specific malignancy risk prediction model in thyroid nodules.

Table 3. Median percentage of correct predictive value of the network following 20 rounds of training and the area under the curve (AUC) for the resilient back propagation (RP) training algorithm.

Median accuracy	ROC analysis			95% confidence interval	
	AUC	Standard deviation	*P*-value	Lower Bound	Upper Bound
64.5%	0.72	0.064	0.003	0.596	0.845

ROC: Receiver operating characteristic curve

We used two different methods to calculate the diagnostic accuracy rate for the trained network - area under the ROC (AUR) and cross-validation. The network trained by the study's clinical dataset utilized the back propagation algorithm with the RP training method. Its diagnostic accuracy rate was 64.66%.

According to Zhu et al.,[10] there were six sonographic features significantly associated with malignant nodules. These features included shape, margin, echogenicity, internal composition, presence of calcifications and peripheral halo. In this study the accuracy of ANN in predicting malignancy of thyroid nodules was 82.3%. The trained network in present study discriminated benign and malignant thyroid nodules with proper accuracy. (Of ANN constructed by sonographic features.)

Ivanov et al.,[11] reported an accuracy of 95.5% for malignant pathology by using ANN.

Saylam et al.,[12] have shown that the ANN technique was of benefit in diagnosing malignancy and might help to reduce unnecessary thyroidectomies in multinodular goiter patients who had indeterminate cytology (FNA) results.

According to this trained network the most important risk factors for malignancy of thyroid nodules were a positive family history of thyroid disease, malignant FNA biopsy result, presence of multiple nodules, disease duration >1.5 years, a positive family history of malignancy, rapid growth of thyroid nodule(s), size of the thyroid nodule (>2 cm), >36 years of age, size of the thyroid gland according to physical examination (>5.5 cm), and gender. Based on these findings, the clinician would be able to decide on a surgical plan by taking into consideration the FNA biopsy results, the number of risk factors noted and their priority.

According to Ippolito et al.,[13] neural network analysis of cytologic data could be a useful tool to refine the cancer risk in patients with lesions diagnosed by fine needle aspiration biopsy (FNAB).

One advantage of modeling by an ANN is its ability to describe the vague and complicated relations between clinical findings without special assumptions about their counts and nature. The sole important point in the modeling process is the dataset size. The larger the data set, the better the network is trained which will increase its diagnostic accuracy rate. Accordingly, the role of each mentioned risk factor in prediction of thyroid nodule malignancy will be better ascertained in a larger study of more data. A comparison among other training algorithms for ANN is recommended for future studies.

Conclusion

Our work strongly agreed with an extensive bio-statistically validated ANN of certain clinical as well as paraclinical and individual given inputs (predictors) that have the capability to stratify the malignancy risk of a thyroid nodule in order to individualize patient care. This risk assessment model (tool) could minimize unnecessary diagnostic thyroid surgeries obtained by the FNA misclassification.

This new, promising risk stratification tool definitely warrants further validation through additional trials. We claim that this model can be easily accessed by physicians via an application (software) and utilized by surgeons to assess the malignancy risk of thyroid nodules in specific subjects.

Conflict of Interest

No conflict of interest is declared.

References

1. Pinchot SN, Al-Wagih H, Schaefer S, Sippel R, Chen H. Accuracy of fine needle aspiration biopsy for predicting neoplasm or carcinoma in thyroid nodules 4 cm or larger. *Arch Surg.* 2009;144(7):649-55.
2. Stang MT, Carty SE. Recent developments in predicting thyroid malignancy. *Curr Opin Oncol.* 2009;21(1):11-7.
3. Wang CC, Friedman L, Kennedy GC, Wang H, Kebebew E, Steward DL, et al. A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid.* 2011;21(3):243-51.
4. Berri RN, Lloyd LR. Defining the role of fine needle aspiration of thyroid nodules in male patients: is it necessary? *Am J Surg.* 2008;195(3):396-8; discussion 8-400.
5. Stojadinovic A, Peoples GE, Libutti SK, Henry LR, Eberhardt J, Howard RS, et al. Development of a clinical decision model for thyroid nodules. *BMC Surg.* 2009;9:12.
6. Tee YY, Lowe AJ, Brand CA, Judson RT. Fine needle aspiration may miss a third of all malignancy in palpable thyroid nodules: a comprehensive literature review. *Ann Surg.* 2007;246(5):714-20.
7. Shang JS, Lin YS, Goetz AM. Diagnosis of MRSA with neural networks and logistic regression approach. *Health Care Manag Sci.* 2000;3(4):287-97.
8. Khan AN, Al-Khattab Y, Macdonald U. Thyroid nodule imaging. [Internet] Available from: http://emedicine.medscape.com/article/385301-overview.
9. Burch HB, Burman KD, Reed HL, Buckner L, Raber T, Ownbey JL. Fine needle aspiration of thyroid nodules. Determinants of insufficiency rate and malignancy yield at thyroidectomy. *Acta Cytol.* 1996;40(6):1176-83.
10. Zhu LC, Ye YL, Luo WH, Su M, Wei HP, Zhang XB, et al. A model to discriminate malignant from benign thyroid nodules using artificial neural network. *PLoS One.* 2013;8(12):e82211.
11. Ivanov SV, Zhabin SN, Gorbacheva OS, Barsukov VS, Denisova EIu. Multifactor analysis of nodular goiter by the method of artificial neural network. *Klin Med (Mosk).* 2008;86(4):49-50.
12. 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(7):554-7.
13. 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(12):1065-71.