

RESEARCH ARTICLE

Editorial Process: Submission:10/31/2017 Acceptance:02/17/2018

Assessing Breast Cancer Risk with an Artificial Neural Network

Mojtaba Sepandi¹, Maryam Taghdir¹, Abbas Rezaianzadeh^{2*},
Salar Rahimikazerooni²

Abstract

Objectives: Radiologists face uncertainty in making decisions based on their judgment of breast cancer risk. Artificial intelligence and machine learning techniques have been widely applied in detection/recognition of cancer. This study aimed to establish a model to aid radiologists in breast cancer risk estimation. This incorporated imaging methods and fine needle aspiration biopsy (FNAB) for cyto-pathological diagnosis. **Methods:** An artificial neural network (ANN) technique was used on a retrospectively collected dataset including mammographic results, risk factors, and clinical findings to accurately predict the probability of breast cancer in individual patients. Area under the receiver-operating characteristic curve (AUC), accuracy, sensitivity, specificity, and positive and negative predictive values were used to evaluate discriminative performance. **Result:** The network incorporating the selected features performed best (AUC = 0.955). Sensitivity and specificity of the ANN were respectively calculated as 0.82 and 0.90. In addition, negative and positive predictive values were respectively computed as 0.90 and 0.80. **Conclusion:** ANN has potential applications as a decision-support tool to help underperforming practitioners to improve the positive predictive value of biopsy recommendations.

Keywords: Breast cancer- artificial neural network- risk assessment

Asian Pac J Cancer Prev, **19** (4), 1017-1019

Introduction

Breast cancer is the commonest cancer in women and one of the most common causes of cancer-related mortality in women worldwide (Ferlay et al., 2010). Proper diagnosis of breast cancer is attained by integration of several clinical variables and mammographic features. An ideal diagnostic system must discriminate between benign and malignant masses (Ayer et al., 2010).

Radiologists must make decisions based on their judgment of breast cancer risk with uncertainty. So there is an interest in developing tools that can calculate an accurate probability of breast cancer to aid radiologists (Elmore et al., 1994). Imaging methods, mostly fine needle aspiration biopsy (FNAB) have been used for establishing cyto-pathological diagnosis of breast cancer (Onur et al., 2015). A study conducted in south of Iran has been shown that the result of about 35% of all FNAB performed on women were benign (Rezaianzadeh et al., 2014). So there is a need to detecting the disease in effectively and accurately (Sountharajan et al., 2017). Artificial intelligence and machine learning techniques have been widely applied in detection/recognition of breast cancer. The ability of diagnostic tool to distinguish between benign and malignant abnormalities is the main component of accuracy in a risk estimation process.

Artificial Neural Network (ANN) technique allow arbitrary nonlinear relations between variables. Standard statistical approaches (e.g., logistic regression) need further modeling processes to allow this flexibility. Furthermore, ANNs do not necessitate distributional assumptions (such as normality). These advantages have caused considerable interest in the use of ANN for medical outcomes (Sargent, 2001). breast cancer risk assessment can improve clinical management of patients (Erbil et al., 2015).

The aim of the present study is to assess whether an ANN trained on a large retrospectively collected dataset of consecutive mammography findings can distinguish between benign and malignant cases and accurately predict the probability of breast cancer for individual patients.

Materials and Methods

The dataset consisted of 655 women (196 malignant and 459 benign) retrieved from 11,850 screened cases referred to Shahid Motahhari breast clinic affiliated to Shiraz University of Medical Sciences between 2004 and 2012. All mammography observations were made by radiologists and all demographic and clinical variables were documented by trained health workers. ANNs

¹Health Research Center, Life style institute, Baqiyatallah University of Medical Sciences, Tehran, ²Colorectal Research Center, Shiraz University of Medical Sciences, Shiraz, Iran. *For Correspondence: msepandi@gmail.com

are computer models consist of highly interconnected nodes, and their overall ability to help predict outcomes is determined by the connections between these neurons. The nodes in different layers are connected with arcs (connection weights). ANNs “learn” the relationships between input variables and the effects they have on outcome by increasing or decreasing the values of these connection weights on the basis of known cases. The procedure of estimating the weights is called learning. A back-propagation learning algorithm was used in this study. There are several supervised training algorithms for ANN. One of the most common is backward propagation of errors (back propagation). This algorithm is based on the error-correction learning rule. The error back-propagation process contains a forward and a backward pass through the different layers of the network. (Chauvin and Rumelhart, 1995).

Malignant and benign cases had been confirmed by pathologic reports. We built our ANN as a three-layer feed-forward network with use of MATLAB 7.4 (Mathworks, Natick, Mass). The layers included an input layer of the 23 variables shown in table 1, a hidden layer with 1000 hidden nodes, and an output layer with a single node. The output node generated a number between 0 and 1 that represented the risk of malignancy.

To train and test the ANN model, ten-fold cross validation was used. In ten-fold cross validation, the data was divided into ten subsets that were approximately equal in size. This process were repeat for ten iterations until all subsets were used once for testing. The dataset is firstly divided into 10 subsets randomly and each time, one of the 10 subsets is used as the test set and the other 9 subsets are used in the training set. Performance of the ANN classification method is evaluated through sensitivity, specificity, and accuracy tests. Sensitivity, specificity, and accuracy are commonly used statistics, using True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP) terms. TP is the number of true positives, which means that some cases with ‘positive’ class are correctly classified as positive. FP refers to the number of false positives, which means that some cases with ‘positive’ class are incorrectly classified as positive and should be in the negative class. TN is the number of true negatives, which means that some cases with ‘negative’ class are correctly classified as negative. Finally, FN refers to the number of false negatives, which means that some cases with ‘negative’ class should be classified as positive. In this study, area under the receiver-operating characteristic curve (AUC), accuracy, sensitivity, specificity, and positive and negative predictive values were used to evaluate discriminative performance of the models. ROC curve is usually used in clinical epidemiology to show how medical diagnostic tests can discriminate between two patient states. The ROC curve demonstrates the trade off between the true positive fraction and false positive fraction as one change the criterion for positivity. An ROC curve lying on the diagonal line reveals the performance of a diagnostic test that is no better than flipping a coin. The AUC is an effective and combined measure of sensitivity and specificity that describes the inherent validity of diagnostic

tests (Hajian-Tilaki, 2013).

Results

The AUC of our ANN, was 0.955 (Figure 1). Sensitivity and specificity of the ANN were respectively calculated as 0.82 and 0.90. Besides, negative and positive predictive values of ANN were respectively computed as 0.90 and 0.80.

Table 1. Variables Used in the ANN

Variables	Levels
Breast Density	Predominantly fatty, scattered fibroglandular, heterogeneously dense, extremely dense
Age Groups, y	<45, 45-50, 51-54, 55-60, 61-64, ≥65
Family History of Breast Cancer	Yes, No
Mass Shape	Circumscribed, ill-defined, microlobulated, spiculated, not present
Mass Margins	Oval, round, lobular, irregular, not present
Mass Density	Fat, low, equal, high, not present
Mass Size	None, small (<3 cm), large (≥3 cm)
Lymph Node	Present, not present
Asymmetric Density	Present, not present
Skin Thickening	Present, not present
Skin Retraction	Present, not present
Nipple Retraction	Present, not present
Skin Lesion	Present, not present
Micro-calcifications	Present, not present
Simpling	Present, not present
History of Breast Surgery	Yes, No
Menopause	Pre-menopause, Postmenopause
Marital Status	Single, Married
history of contraceptive use	Yes, No
age at first pregnancy	<30y, ≥30y
occupation	Housewife, Employee
parity	0, 1, 2, 3, ≥4
age at menarche	<12y, ≥12y

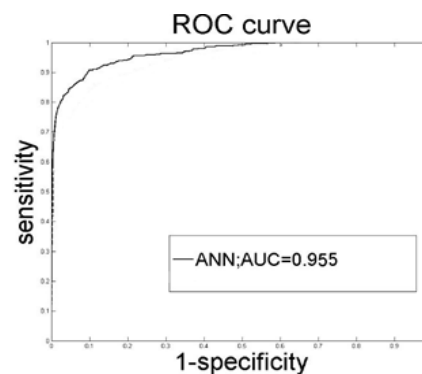


Figure 1. ROC Curve Created from the Output Probabilities of Our ANN. AUC, Area under the ROC curve.

Discussion

The results have showed that our ANN can precisely estimate the risk of breast cancer by using a dataset that contains demographic data and mammographic features. Similar to other ANN models presented in the literature, our ANN has the potential to aid radiologists in classifying patients. When compared with the previous models developed by our research team (a Bayesian network), the discrimination performance of our ANN was slightly higher (ANN AUC, 0.955; Bayesian network AUC, 0.940) (Rezaianzadeh et al., 2016). Our results must be viewed with some caution with respect to their generalizability because significant variability has been observed in the interpretive performance of screening and diagnostic mammography. In the present study, an ANN breast cancer risk estimation model was constructed based on 1,1850 screened cases referred to Shahid Motahhari breast clinic affiliated to Shiraz University of Medical Sciences between 2004 and 2012 to aid physicians in breast cancer diagnosis. The results demonstrated that the ANN could perform well in estimating the probability of malignancy, and improve the positive predictive value (PPV) of the decision to perform biopsy. Using a computer-assisted detection program as a second reader has been shown to improve sensitivity in the screening setting (Freer and Ulisse, 2001). However, given our present results, our ANN had the potential to be used as a decision-support tool that could help underperforming practitioners improve the PPV of biopsy recommendations. Our model reinforced the previously known mammography predictors of breast cancer; i.e., irregular mass shape, speculated mass margins, micro calcifications, and breast density (Chhatwal et al., 2009).

One of the limitations of this study was a retrospective analysis using registered data. This limit the external validity of the results, thus, future testing in a larger population is recommended. Breast Imaging-Reporting and Data System (BI-RADS) categories were not assessed in the present study, due to lack of the technology in our center. Adding BI-RADS, can improve the model's performance (Lo et al., 2002). In conclusion the authors' ANN could effectively discriminate benign masses from malignant ones.

References

- Ayer T, Alagoz O, Chhatwal J, et al (2010). Breast cancer risk estimation with artificial neural networks revisited. *Cancer*, **116**, 3310-21.
- Chauvin Y, Rumelhart DE (1995). Backpropagation: theory, architectures, and applications. Psychology press, pp 220-2.
- Chhatwal J, Alagoz O, Lindstrom MJ, et al (2009). A logistic regression model based on the national mammography database format to aid breast cancer diagnosis. *AJR Am J Roentgenol*, **192**, 1117-27.
- Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR (1994). Variability in radiologists' interpretations of mammograms. *N Engl J Med*, **331**, 1493-9.
- Erbil N, Nursel D, Çiğdem İ, Bölükbaş N (2015). Breast cancer risk assessment using the Gail model: a Turkish study. *Asian Pac J Cancer Prev*, **16**, 303-6.

- Ferlay J, Héry C, Autier P, Sankaranarayanan R (2010). Global burden of breast cancer. *Breast cancer epidemiology*: Springer, pp 1-19.
- Freer TW, Ulisse MJ (2001). Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology*, **220**, 781-6.
- Hajian-Tilaki K (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med*, **4**, 627.
- Lo JY, Markey MK, Baker JA, Floyd Jr CE (2002). Cross-institutional evaluation of BI-RADS predictive model for mammographic diagnosis of breast cancer. *AJR Am J Roentgenol*, **178**, 457-63.
- Onur GO, Tarcan E, Onur A, et al (2015). Comparison between radiological and invasive diagnostic modalities in diagnosis of breast cancer. *Asian Pac J Cancer Prev*, **16**, 4323-8.
- Rezaianzadeh A, Sepandi M, Akrami M, et al (2014). Pathological profile of patients with breast diseases in Shiraz. *Asian Pac J Cancer Prev*, **15**, 8191.
- Rezaianzadeh A, Sepandi M, Rahimikazerooni S (2016). Assessment of breast cancer risk in an Iranian female population using bayesian networks with varying node number. *Asian Pac J Cancer Prev*, **17**, 4913.
- Sargent DJ (2001). Comparison of artificial neural networks with other statistical approaches. *Cancer*, **91**, 1636-42.
- Sountharajan S, Karthiga M, Suganya E, Rajan C (2017). Automatic classification on bio medical prognosis of invasive breast cancer. *Asian Pac J Cancer Prev*, **18**, 2541-4.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.