

RESEARCH ARTICLE

Editorial Process: Submission:09/04/2017 Acceptance:12/30/2017

Predicting the Survival of Gastric Cancer Patients Using Artificial and Bayesian Neural Networks

Azam Korhani Kangi, Abbas Bahrampour*

Abstract

Introduction and purpose: In recent years the use of neural networks without any premises for investigation of prognosis in analyzing survival data has increased. Artificial neural networks (ANN) use small processors with a continuous network to solve problems inspired by the human brain. Bayesian neural networks (BNN) constitute a neural-based approach to modeling and non-linearization of complex issues using special algorithms and statistical methods. Gastric cancer incidence is the first and third ranking for men and women in Iran, respectively. The aim of the present study was to assess the value of an artificial neural network and a Bayesian neural network for modeling and predicting of probability of gastric cancer patient death. **Materials and Methods:** In this study, we used information on 339 patients aged from 20 to 90 years old with positive gastric cancer, referred to Afzalipoor and Shahid Bahonar Hospitals in Kerman City from 2001 to 2015. The three layers perceptron neural network (ANN) and the Bayesian neural network (BNN) were used for predicting the probability of mortality using the available data. To investigate differences between the models, sensitivity, specificity, accuracy and the area under receiver operating characteristic curves (AUROCs) were generated. **Results:** In this study, the sensitivity and specificity of the artificial neural network and Bayesian neural network models were 0.882, 0.903 and 0.954, 0.909, respectively. Prediction accuracy and the area under curve ROC for the two models were 0.891, 0.944 and 0.935, 0.961. The age at diagnosis of gastric cancer was most important for predicting survival, followed by tumor grade, morphology, gender, smoking history, opium consumption, receiving chemotherapy, presence of metastasis, tumor stage, receiving radiotherapy, and being resident in a village. **Conclusion:** The findings of the present study indicated that the Bayesian neural network is preferable to an artificial neural network for predicting survival of gastric cancer patients in Iran.

Keywords: Survival- gastric cancer- Bayesian neural networks- artificial neural network

Asian Pac J Cancer Prev, 19 (2), 487-490

Introduction

According to annual recording of 1,000,000 cases, gastric cancer is currently the prevalent for gastrointestinal cancer (Plummer et al., 2016) and gastric cancer is the 3rd leading cause of death globally (Rugge et al., 2015). Gastric cancer showed first and third ranking for men and women in Iran, respectively (Veisani and Delpisheh, 2016). Because the signs of gastric cancer is normal and is like to digestive diseases, normally prognosticating the Sickness is done at the advanced level of the sickness. In this level all of the treatment methods has no significant effect on the survival of the patients (Conteduca et al., 2013; Malekzadeh et al., 2009). According to the report of the world health organization (WHO), in 2012 gastric cancer was the 3rd leading cause of death globally between two genders and the first leading cause of death by 15.5% of the Total. Nowadays by development of technology and complication of the issues, new methods were replaced by traditional ones that had no capability

to predict the situation properly. It should be added that complicating the issues leads to problems such as non-linearization equation between produced parameters that traditional methods have no capability to perform and despite of high longevity and easiness of use, error rate of their prognosticating (predicting) is high and have no considerable accuracy (Abdel-Aal, 2006). Considering the premises such as linearization equation between dependent and independent variables, consistency of error variance, lack of linearization equation and sensitivity of these models to outlier data, we have some limitations, so the new methods such as neural network are used to analyze these issues. These networks Process information exactly as the human brain's processing methods. This method is applicable in problem solving of the methods that have no clearance equation between variable, because without defining the equation between each variable, it can learn the variable equation. Neural networks are composed of many neurons that work together to solve a certain problem and neurons have to exchange information with

each other and each neuron has separate weight. In other word, adjusting the input weights of each neuron cause the learning of the whole network. Each neural network is formed by three layers including: input layer that receive information, many hidden layers that receives information from the previous layer and finally the output layer that the results of calculation and responses are placed in it. Learning algorithm is a process that adjusts the network weight. The purpose of the training network is to teach law of work to network in a way that after learning is able to produce the right output for each input (Norusis, 1994). Network structures are designed as a way to describe neuron connection and estimate network parameters (including weight and bias) by using suitable learning algorithm. Multi-layers perceptron neural networks by minimizing output mean square error and learning algorithm and poster observer-error and is one prevalent to artificial neural network and compared to another networks structures (Hopfield, RBF...) is powerful calculator in prediction and categorization (Kay and Titterington, 1999).

input of each neuron is obtained from the total weight of previous attached output and then function F comes into this set, so

$$net = \sum_j w_{ij} \times output_j + w_i$$

W_i is the bias of i and reflexes input impact rate on neuron. In this model each connection has a separate weight that can be shown by W_{ij} and would be considered as the rate for connection power of j to i .

Using the Bayesian algorithm is another method for adjusting optimal parameters in neural network that automatically adjust suitable rate for function parameters (MacKay, 1992). There are many algorithms for weighting and training of neural network which of them the new method is Bayesian algorithm. In Bayesian method, output of the test mode can be obtained from all possible values of neural networks parameters, in other word we have one output for each selection. In Bayesian method one weight class is not efficient but one distribution function would be considered for w and In Bayesian method one weight class is not efficient but one distribution function would be considered for w and all affected weighting to be considered for calculating final output. This method can reduce extra matches (Bishop, 1997). Bayesian method can be calculated by the following equation

$$F(x) = \beta E_D + \alpha E_w$$

According to the above equation, E_D is residual sum of squares and E_w is weight sum of squares.

α and β are objective function parameters which are calculated by Gauss-Newton method (MacKay, 1992).

Although Bayesian neural network are less used in medical issues, but many researches had been done to predict the results of cancers by artificial neural networks and compare these models to old Statistical ones and in regard the following subjects can be mentioned.

Zhu et al., (2013) used artificial neural networks and Cox regression models to predict the survival of gastric

cancer patients. The findings of their research indicated that neural network model accuracy of 85.3% and the area under receiver operating characteristic curve (AUROC) by accuracy of 89% had a better performance than Cox regression by accuracy of 81.9% and Load0deformation curve by accuracy of 82.4%. Biglerian et al., (2010) used an artificial neural network and Viable parametric regression models to predict the survival of gastric cancer patients. The result of their study showed that prognosis level and AUROC for two artificial neural networks and viable models are (79.45% and 81.5%) in comparison to (73.97% and 74.8%), respectively. Their finding indicated that prediction of artificial neural network was better than Viable parametric model Wang et al., (2009) employed neural network and decision Tree to predict hospital cost of gastric cancer patients and their findings showed that absolute error and correlation coefficient of two artificial neural network and decision tree was (116,279 in comparison to 3,424,908) and (0.987 in comparison to 0.806), respectively and the result of their research showed that artificial neural network had a better performance than decision tree to predict hospital costs of gastric cancer patients. Hajizadeh et al., (2010) applied two artificial neural network and Cox regression to predict the survival of the gastric cancer patients. Their findings showed that Prognosis accuracy and AUROC for two artificial neural network and Cox regression models were 81.5, 82.6 in comparison to 72.6 and 75.4, respectively. The result of their research indicated that artificial neural network had a better performance than Cox regression model. Khan et al., (2005) used Bayesian neural network to predict rainfall-runoff Sagoni pond model in Canada. Their finding regarding to the criteria such as R^2 , RMSE, PFC, LFC showed that Bayesian neural network by precision of (0.97, 6.45, 0.09, 0.14) had a high precision and minimal error in predicting runoff, than artificial neural network by accuracy of (0.95, 8.84, 0.10, 0.15) (Khan and Coulibaly, 2006). Fithriasari et al., (2013) employed Bayesian neural network to estimate hourly rainfall. The finding of their study showed that total square Error for Bayesian neural network and feed forward neural network was 2.16 and 2.52, respectively.

The present study aims to analyze the survival of gastric cancer patients and predict the mortality of these patients by two Bayesian neural network and artificial neural network.

Materials and Methods

In this retrospective study 339 patients with positive gastric cancer were studied. Medical records of the patients who were referred to Bahonar and Afzalipoor Hospitals during the year 2001 to 2015 were examined and after registration, patient names were matched with cancer record data related to Kerman City and related mortality data. Used independent variables in this research are age, gender, smoking, addiction to opium, place of living (rural and urban district), surgical history, chemotherapy history, radiotherapy history, histological grade, cancer staging, Morphology, Metastasis, family history, longevity after dieses diagnosis up to research date. Used dependent

variables in this research are: patient's survival status that has been coding as Two-state variable with codes zero (0) for death and one (1) for survival. In this study artificial neural network and bayesian neural network were both used to analyze data. In the first step of the modeling process, data were divided in training (70% of patients) and testing (30% of patients) subsets and then the model fitted to all the sample size. To compare subgroups, we used the log-rank test. In the case of education and selecting the proper architecture for network, Models with input parameters were used and some different neuron in hidden layer was processed. Used artificial neural network in the education process was a three layer back propagation neural network by 30 neurons in input layer, 9 neurons in hidden layer and 2 neurons in output layer by using learning back-propagation algorithm and sigmoid activation function with learning rate of 0.01 to 0.4 and momentum 0.8 to 0.95 were. By using approximate Gauss-newton method for adjusting input parameters, for selecting the best structure for Bayesian neural network, the following selection were made: network with 3 hidden layers, selecting middle layer nodes by trial and error method, selecting network with 14 nodes in input layer, selecting 5 nodes in hidden layer, selecting 1 node in output layer. Sensitivity, specificity, accuracy and AUROC are used to compare the prediction of two models artificial neural network and Bayesian neural network. For analyzing the data, Matlab2011a and Spss 18 software are applicable.

Results

The finding of the study indicated that among 339 patients, 195 patients (57.5%) had passed away and 144 patients (42.5 %) were alive or we were not aware of their survival status. The number of men and women patients was 216 (63.7) and 123 (36.3), respectively. Average survival rate for patients was 21.70 months. Symptoms of the disease in men and women were recognized in the age of 60-70 years old and 70-80 years old, respectively. Table 1 shows real and predicted data by two models. Table 2 shows a descriptive information related to dependent variables. Table 3 shows ANN modeling results of prognostic factors on gastric cancer patient survival. In Table 4 comparison indicators of two models are presented. Regarding to obtained real and predicted data by two artificial and Bayesian networks (Table 1) and

Table 1. Observations and Predicted Cases by Two Artificial Neural Network Models and Bayesian Neural Network

predicted	observations	
	Death	Survival
Artificial neural network		
Death	172	14
Survival	23	130
Bayesian neural network		
Death	186	13
Survival	9	131

Table 2. Independent Variables Used in Artificial Neural Network Model and Bayesian Neural Network

Variable Name	No. (%)	Variable Name	No. (%)
Mean± standard deviation from the time of diagnosis	21.70±20.38	Surgery	
Mean± standard deviation of age	62.84±14.52	Yes	274
Gender		No	65
Male	216	Chemotherapy	
Female	123	Yes	135
Smoking		No	204
Yes	91	Radiotherapy	
No	248	Yes	46
Opium		No	293
Yes	124	Metastasis	
No	215	Yes	134
Rural		No	205
Yes	67	Histological grade	
No	272	G1	11
Family history		G2	270
Yes	29	G3	58
No	310	Cancer Staging	
Morphology		I	9
Neoplasm	23	II	205
Adenocarcinoma	263	III	91
Carcinoma	53	IV	34

criteria rate for comparison of two models (Table 4), it can be concluded that in all compared criteria, Bayesian neural network had a perfect proficiency over artificial neural network.

Table 3. ANN Modeling Results of Prognostic Factors on Gastric Cancer Patient Survival

Ordered factors	Normalized importance
Age	0.664
Histological grade	0.37
Morphology	0.344
Sex	0.326
Smoking	0.322
Opium	0.297
Chemotherapy	0.272
Metastasis	0.252
Cancer Staging	0.246
Radiotherapy	0.228
Rural	0.21
Family history	0.198
Surgery	0.141

Table 4. Comparative Performance Indices of ANN and BNN Models

model	Sensitivity	Specificity	Accuracy	AUROC
artificial neural network	0.882	0.903	0.891	0.944
bayesian neural network	0.954	0.909	0.935	0.961

Discussion

Cancer is one of the important health problems in the world. In Iran, gastric cancer is the most lethal cancer and the first common cancer in men. Because of delayed diagnosis of the cancer and it being in advanced level, survival of the gastric cancer patients is short. Many studies had been done to analyze the survival of the gastric cancer patient and in this research, we aim to analyze data by using artificial neural network and Bayesian neural network. In the recent years the use of neural network without any premises for predicting of model and prognosis has increased, in this regard many studies had been done in different countries.

According to the study that had been done by Zhu et al., (2013) to predict the survival of gastric cancer patients by using two artificial neural networks and Cox regression, it was found that accuracy of artificial neural network was higher than Cox regression. Wang et al., (2009) done in China, a study by using neural network and decision tree to predict the hospital costs of gastric cancer patients. Their finding indicated that artificial neural network regarding to predicting of hospital cost of gastric cancer patients, had more efficiently than Decision tree.

In the present study we used artificial neural network and Bayesian neural network top predict the survival of gastric cancer patients and the following finding was obtained.

Sensitivity and specificity of artificial neural network and Bayesian neural network were (0.882, 0.903) and (0.954, 0.909), respectively. Accuracy and AUROC were (0.891, 0.944) and (0.935, 0.961), respectively. The findings of this study showed that due to high sensitivity indicator, high specificity, high accuracy and the area under receiver operating characteristic curve (AUROC) Bayesian neural network is a preferable model than artificial neural network for predicting and categorizing the survival data.

In conclusion, the findings of this study showed that for predicting the survival rate of gastric cancer, Bayesian neural network is preferable method over artificial neural network and diagnosis at an early age and in primary levels of sickness can long the survival of the patients.

References

- Abdel-aal R (2006). Modeling and forecasting electric daily peak loads using abductive networks. *Int J Electr Power Energy Syst*, **28**, 133-41.
- Biglarian A, Hajizadeh E, Kazemnejad A (2010). Comparison of artificial neural network and Cox regression models in survival prediction of gastric cancer patients. *Koomesh*, **11**, 215-20.
- Bishop CM, Christopher M (1997). Neural Networks. *J Braz Comp Soc Campinas*, **4**, 1, Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-65001997000200006&lng=en&nrm=iso>. access on 04 Jan. 2018. <http://dx.doi.org/10.1590/S0104-65001997000200006>.
- Conteduca V, Sansonno D, Lauletta G, et al (2013). H. pylori infection and gastric cancer: state of the art. *Int J Oncol*,

42, 5-18.

- Fithriasari K, Iriawan N, Ulama BSS, Kuswanto H (2013). Prediction of hourly rainfall using Bayesian neural network with adjusting procedure. The third basic science international conference.
- Hajizadeh E, Kazemnejad A, Biglarian A (2010). Comparison of artificial neural network and Cox regression models in survival prediction of gastric cancer patients.
- Kay JW, Titterton DM (1999). Statistics and neural networks: advances at the interface, Oxford University Press on Demand.
- Khan MS, Couibaly P (2006). Bayesian neural network for rainfall-runoff modeling. *Water Resour Res*, **42**, 1-18.
- Mackay DJ (1992). Bayesian interpolation. *Neural Comput*, **4**, 415-47.
- Malekzadeh R, Derakhshan MH, Malekzadeh Z (2009). Gastric cancer in Iran: epidemiology and risk factors. *Arch Iran Med*, **12**, 576-83.
- Norusis MJ (1994). SPSS, SPSS Incorporated.
- Plummer M, DE Martel C, Vignat J, et al (2016). Global burden of cancers attributable to infections in 2012: a synthetic analysis. *Lancet Glob Health*, **4**, 609-16.
- Rugge M, Fassan M, Graham DY (2015). Epidemiology of gastric cancer. Gastric cancer. Springer, pp 23-8.
- Veisani Y, Delpisheh A (2016). Survival rate of gastric cancer in Iran; a systematic review and meta-analysis. *Gastroenterol Hepatol Bed Bench*, **9**, 78.
- Wang J, Li M, Hu Y-T, Zhu Y (2009). Comparison of hospital charge prediction models for gastric cancer patients: neural network vs. decision tree models. *BMC Health Serv Res*, **9**, 161.
- Zhu L, Luo W, Su M, et al (2013). Comparison between artificial neural network and Cox regression model in predicting the survival rate of gastric cancer patients. *Biomed Rep*, **1**, 757-60.



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