International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:7, No:7, 2013

Comparison of Neural Network and Logistic Regression Methods to Predict Xerostomia after Radiotherapy

Hui-Min Ting, Tsair-Fwu Lee, Ming-Yuan Cho, Pei-Ju Chao, Chun-Ming Chang, Long-Chang Chen, and Fu-Min Fang

Abstract—To evaluate the ability to predict xerostomia after radiotherapy, we constructed and compared neural network and logistic regression models. In this study, 61 patients who completed a questionnaire about their quality of life (QoL) before and after a full course of radiation therapy were included. Based on this questionnaire, some statistical data about the condition of the patients' salivary glands were obtained, and these subjects were included as the inputs of the neural network and logistic regression models in order to predict the probability of xerostomia. Seven variables were then selected from the statistical data according to Cramer's V and point-biserial correlation values and were trained by each model to obtain the respective outputs which were 0.88 and 0.89 for AUC, 9.20 and 7.65 for SSE, and 13.7% and 19.0% for MAPE, respectively. These parameters demonstrate that both neural network and logistic regression methods are effective for predicting conditions of parotid glands.

Keywords**—**NPC, ANN, logistic regression, xerostomia.

I. Introduction

NASOPHARYNGEAL carcinoma (NPC) cancer is a endemic disease in southern China and Taiwan. NPC is highly sensitive to ionizing radiation, and the primary treatment for NPC is radiotherapy [1], [2] Radiation-induced xerostomia is one of the complications resulting from radiotherapy [3], [4]. Xerostomia, which is caused by the damage to the parotid glands, causes difficulties in swallowing and dental problems [5].

In this study, some parameters that can estimate how radiation affects salivary glands are chosen by a statistical method in order to compare the capability of a neural network with a logistic regression model. These parameters are then adopted to build models. The receiver operating characteristic (ROC), sum of squares error (SSE) and mean absolute percentage error (MAPE) are used to evaluate which method is better in terms of predicting xerostomia.

Hui-Min Ting, Tsair-Fwu Lee, Ming-Yuan Cho, Pei-Ju Chao, Chun-Ming Chang, and Long-Chang Chen are with the National Kaohsiung University of Applied Sciences (Corresponding authors: phone: 886-7-3814526; fax: 886-7-3811186; e-mail: tflee@kuas.edu.tw, fang2569@gmail.com).

Hui-Min Ting, Pei-Ju Chao, and Fu-Min Fang are with the Department of Radiation Oncology, Chang Gung Memorial Hospital-Kaohsiung Medical Center, Chang Gung University College of Medicine (e-mail: hmting618@gmail.com).

Long-Chang Chen is with the Department of Radiation Oncology, Chang Gung Memorial Hospital-Chiayi Medical Center, Chang Gung University College of Medicine.

II. Methods and Materials

A. Patient Characteristics

A total of 61 patients who were treated with intensity modulated radiation therapy (IMRT) completed a questionnaire about quality of life (QoL), including personal profile, diagnosis, treatment doses, and cancer stage. The characteristics of the patients are listed in Table I. Patients with xerostomia before treatment was excluded. The remaining patients were categorized into two groups (A and B), corresponding to with and without xerostomia after radiotherapy. (Institutional Review Board No. 99-1420B, 96-1231B)

B. Statistic Methods

1. Cramer's V Correlation

According to the theory of categorical data analysis, observed and expected frequencies are calculated to obtain a chi-squared value. A smaller chi-squared value means that the observed frequency is closer to the expected frequency. However, a chi-squared value does not give any information about how one event is related to another event. The correlation coefficient is introduced to address this problem. The correlation coefficient varies from 0 to 1. The linear relationship between events is strong if the coefficient is close to 1; otherwise, linear relationship is weak.

Here, Cramer's V is adopted to select 10 categorized variables of xerostomia. The variables include cancer stage (e.g. TNM stage and AJCC stage), and patient profile (e.g. education, gender).

2. Point-Biserial Correlation

A continuous variable may have uncountable outcomes which vary in a certain range and with a specific unit. Assume variable X is a continuous variable and variable Y is dichotomous. The correlation between variables X and Y is called point-biserial correlation. In this study, point-biserial correlation is used to select seven continuous variables (e.g. parotid mean dose).

TABLE I
PATIENTS AND TUMOR CHARACTERISTICS

Characteristics	Catagory	Npc (N=61)%
Gender	Male	77.05
	Female	22.95
Education	Illiteracy	1.64
	Primary	11.48
	Junior+Senior	49.18
	College Or Above	37.70
Age	< 39	22.95
	40-49	32.79
	50-59	32.79
	60-69	8.20
	70-79	3.28
Chemotherapy	Yes	1.64
	No	98.36
Parotid Dose	Mean (Gy)	41.02±13.33 (Gy)
Ajcc Stage	1	3.28
	2	19.03
	3	34.43
	4	24.59
	5	19.67
Tnm Stage	T1	34.43
	T2	3.28
	T3	19.67
	T4	14.75
	T5	8.20
	T6	19.67
	N0	18.03
	N1	40.98
	N2	31.15
	N3	0.00
	N4	3.28
	N5	6.56

Abbreviation: AJCC: American Joint Committee on Cancer

C. Models

Variables chosen by Cramer's V and point-biserial correlation are analyzed to build neural network and logistic regression models.

1. Logistic Regression

Logistic regression is a qualitative dependent variable regression model and capable of handling dichotomous problems. The model estimates the chance of suffering from xerostomia and presents the outcome as a probability whose value is restricted to between 0 and 1, with a threshold value of 0.5. If the probability is greater than 0.5 then the patient has xerostomia, otherwise they do not. The model is described in (1):

$$\ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1\chi_1 + \ldots + \beta_i\chi_i \tag{1}$$

where P_i denotes the probability of xerostomia, χ_i denotes the predictive parameters and β_i denotes the parametric coefficients [6].

2. Neural Network

Leave-one-out cross validation (LOOCV) is the training method adopted in the network. Each of N sets of data is taken as input in turn and one is trained by the others (N-1 sets). N iterations are required in the whole training process. This method is ideal when the amount of data is small and high accuracy is needed [7].

The learning model adopted in the planned network is so-called pattern recognition which is a type of feed-forward neural network. Neurons of each layer only receive signals from the previous layer, as shown in Fig. 1 [8], [9]. The fundamental structure of a feed-forward neural network is described as follows:

1. *Input Layer*: This layer is in charge of receiving variables and it can be classified into two types. In the first, neurons in the input layer possess transfer functions, weights, and biases which mean that their outputs are delivered to the next stage after operations, whereas in the second, variables just pass through neurons without any calculations. The number of neurons in this layer depends on situations; the first type of input layer with operation ability is adopted in this study.
2. *Hidden Layer*: An interface between the input and output layer is called the hidden layer which receives signals from the input layer and operates with its transfer function. The calculated results of neurons are then transmitted to an output layer. It should be mentioned that the hidden layer is not always necessary for a planned structure; no layer, a single layer or multiple layers can be used and this is adjustable by designers.
3. *Output Layer*: A last layer is the output of neurons and the number of neurons is equal to the number of output variables needed.

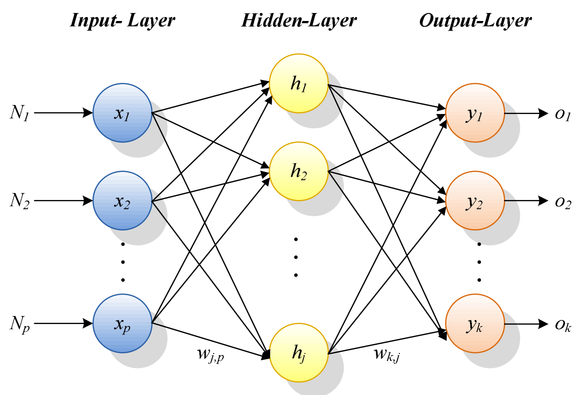

Fig. 1 A diagram of feed-forward neural network

D. Assessment for Models

1. Receiver Operating Characteristic (ROC)

The planned models are built by both a logistic regression and a neural network. Their output values are depicted as ROC curves [10]. It is convenient for us to compare the areas under the ROC curves with each other and identify the better model.

The discrimination is better when the area under the ROC curve is close to 1.

2. Sum of Squares Error (SSE)

This parameter indicates how a prediction value agrees with an observation value. The smaller the value, the less discrete degree they do. It also indicates better prediction ability. Equation (2) shows the formula of SSE [11]:

$$SSE = \sum (y - y')^2 \quad (2)$$

where y denotes the observed value and y' denotes the prediction value.

3. Mean Absolute Percentage Error (MAPE)

The MAPE shows the error between a prediction value and an observation value as a percentage. A smaller MAPE value means that the prediction result is closer to the expected result. The MAPE is described in (3) and it is classified into four levels, listed in Table II [11]:

$$MAPE = \frac{1}{n} \sum \left| \frac{y - y'}{y} \right| \times 100\% \quad (3)$$

where y denotes the observed value, y' denotes prediction value and n denotes the number of prediction values.

TABLE II
CRITERIA OF MAPE

MAPE	Prediction results
<10%	Excellent
10%~20%	Good
20%~50%	Acceptable
>50%	Incorrect

Abbreviation: MAPE: Mean absolute percentage error

III. RESULTS

According to the theory of Cramer's V, categorized variables such as N-stage, T-stage, financial condition and condition of xerostomia before treatment were chosen. Continuous variables, such as gland mean dose, were chosen by point-biserial correlation. All of the variables mentioned above were included in both a logistic regression and a neural network model to obtain ROC values, which were 0.89 and 0.88 as shown in Figs. 2 and 3, respectively.

The SSE was used to evaluate the outcome of the two models. The neural network model score was 9.207 and the logistic regression score was 7.655. Evaluating the outcomes by MAPE, a neural network gave a value of 13.77% compared with a logistic regression value of 19.03%. According to these scores, both neural network and logistic regression are expected to be good methods to predict xerostomia.

Three evaluation methods mentioned above, ROC, SSE and MAPE, were used for a performance test; test results are listed in Table III. As shown in the table, we can see that a logistic regression model is better than a neural network by AUC and SSE evaluation; however, according to the MAPE method, a neural network is superior to logistic regression.

TABLE III
PERFORMANCE OF THREE PREDICTION MODELS

Prediction Model	AUC	SSE	MAPE
Logistic Regression	0.89	7.655	19.03%
Neural Network	0.88	9.207	13.77%

Abbreviation: AUC: the area under the receiver operating characteristic curve;
SSE: sum of squares error; MAPE: mean absolute percentage error

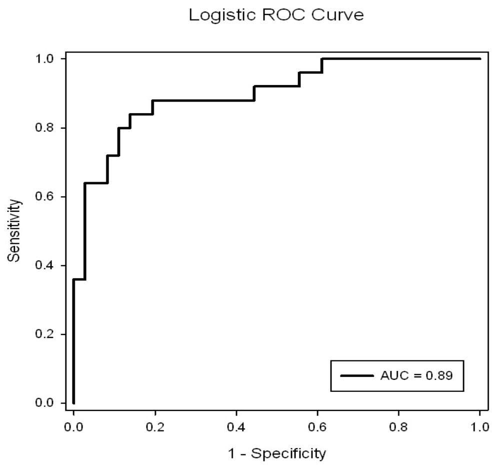

Fig. 2 The receiver operating characteristic curve (ROC) and AUC of logistic regression

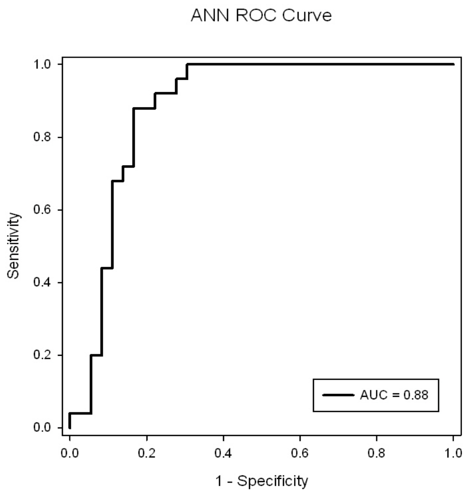

Fig. 3 The receiver operating characteristic curve (ROC) and AUC of neural network (ANN)

IV. DISCUSSION

There are many prediction parameters involved in this study for a single purpose, which is to train prediction models. However, results show that not only gland dose but also N-stage and T-stage parameters affect how prediction models work. Based on the research of Liu M. Z., et al. [12], 749 patients with head and neck cancer accepted chemotherapy and radiation therapy. The cancer staging system is also used to categorize patients in order to improve accuracy of staging. Our staging system includes N-stage, T-stage and AJCC-stage parameters to define the stage of a tumor, but we want to avoid a collinear issue which may arise if the three staging parameters are highly correlated with each other. Performing a collinear analysis ensures that a collinear problem will not occur in our models; in other words, the three staging parameters remain in order to ensure that the prediction models keep functioning well.

The research by Lin A., et al. [13] of QoL of patients with head and neck cancer who took radiation therapy for two years found prediction factors in the EORTC QLQ-C30 questionnaire to include family income, tumor position, AJCC staging, treatment methods and radiotherapy techniques; prediction factors in the EORTC H&N35 questionnaires included tumor position and radiotherapy techniques. Fang F. M., et al. [14] investigated patients' life quality with nasopharyngeal carcinoma cancer and without recurrence for two years. This research suggested that patients have a better QoL if they possess either high education or high income and acceptance of advanced radiotherapy techniques. Because a highly educated patient is usually younger than the average age, this usually means that his or her parotid glands will get better sooner. Also, patients with a good financial position could afford to buy fine food and nutrients which are helpful in the recovery of the salivary glands and for remaining in a good condition.

Beetz I. et al. [15] found that that mean dose of the parotid glands and the condition of salivary glands are major parameters in normal tissue complication probability (NTCP) models for patient-rated xerostomia and sticky saliva and this is the reason why these parameters were included in our models for the assessment of xerostomia.

A theory proposed by Borque A., et al. [16] suggests that neural networks and logistic regression analyses should be used to predict the pathological stage of patients with radical prostatectomy. This research indicated that the accuracy of a neural network was 88.2% compared with 84% for logistic regression, and concluded that the former method was better. In Song's models [17] for computer-aid diagnosis of breast masses, the scores of a neural network and a logistic regression were 85.6% and 85.3% respectively. Eftekhar B., et al. [18] adopted neural networks and logistic regression analysis to predict mortality of head trauma, and showed the AUC for each model to be 96.46% and 95.38% respectively. However, as shown in the three published studies mentioned above, prediction models do not show much difference in score; in other words, neural networks, in addition to logistic regression analysis, are capable of being prediction models for a specific purpose, e.g. in a biomedical field. Three assessment methods were adopted in this study, namely AUC, SSE and MAPE. The AUC method gave values of 0.88 for a neural network model and 0.89 for logistic regression. The evaluated result of SSE suggested that a logistic regression model is better, but this was not true in the case of MAPE. The reason is that while the squared operation in SSE makes all errors positive, the errors are also amplified or reduced. This is different from MAPE, in which only the average errors are taken into consideration. However, these inconsistent results do not mean that our proposed models are unreliable because, as mentioned previously, the scores of each model are very close to each other.

V. Conclusion

In this research, logistic regression models and neural networks for the prediction of xerostomia are compared, and these models give very similar results. A logistic regression model is the preferred model for prediction of xerostomia in nasopharyngeal carcinoma cases after treatment because a neural network is more complex to achieve.

Acknowledgment

This study was supported financially, in part, by grants from the National Science Council (NSC) of the Executive Yuan of the Republic of China. (NSC-101-2221-E-151-007-MY3)

References

[1] T. F. Lee, P. J. Chao, H. M. Ting, S. H. Lo, Y. W. Wang, C. C. Tuan, F. M. Fang, and T. J. Su, "Comparative analysis of SmartArc-based dual arc volumetric-modulated arc radiotherapy (VMAT) versus intensity-modulated radiotherapy (IMRT) for nasopharyngeal carcinoma," *J Appl Clin Med Phys,* vol. 12, p. 3587, 2011.

[2] C. Y. Hsiung, H. M. Ting, H. Y. Huang, C. H. Lee, E. Y. Huang, and H. C. Hsu, "Parotid-sparing intensity-modulated radiotherapy (IMRT) for nasopharyngeal carcinoma: preserved parotid function after IMRT on quantitative salivary scintigraphy, and comparison with historical data after conventional radiotherapy," *Int J Radiat Oncol Biol Phys,* vol. 66, pp. 454-61, 2006.

[3] M. Agulnik and J. B. Epstein, "Nasopharyngeal carcinoma: current management, future directions and dental implications," *Oral Oncol,* vol. 44, pp. 617-27, 2008.

[4] Y. P. Talmi, Z. Horowitz, L. Bedrin, M. Wolf, G. Chaushu, J. Kronenberg, and M. R. Pfeffer, "Quality of life of nasopharyngeal carcinoma patients," *Cancer,* vol. 94, pp. 1012-7, 2002.

[5] V. W. Wu, M. T. Ying, and D. L. Kwong, "Evaluation of radiation-induced changes to parotid glands following conventional radiotherapy in patients with nasopharygneal carcinoma," *Br J Radiol,* vol. 84, pp. 843-9, 2011.

[6] J.-C. Shyu and H.-Y. Liou, " The financial distress prediction model under consideration of business cycle and industry factors - the application of logistic regression model and DEA-DA model " Journal of Risk Management, vol. 12, pp. 157-183, 2010.

[7] J. O. Deasy, J. R. Alaly, and K. Zakaryan, "Obstacles and advances in intensity-modulated radiation therapy treatment planning," Front Radiat Ther Oncol, vol. 40, pp. 42-58, 2007.

[8] M. Isaksson, J. Jalden, and M. J. Murphy, "On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications," *Med Phys,* vol. 32, pp. 3801-9, 2005.

[9] L. Zhang, L. Jia, and W. Zhu, "Overview of traffic flow hybrid ANN forecasting algorithm study," in *2010 International Conference on Computer Application and System Modeling (ICCASM),* , 2010, pp. V1-615-V1-619.

[10] M. H. Zweig and G. Campbell, "Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine," *Clin Chem,* vol. 39, pp. 561-77, 1993.

[11] H.-L. Chen and D.-W. Tsai, "A study of predictive abilities for different models," *Journal of Soil and Water Conservation,* vol. 37, pp. 127-138, 2005.

[12] M. Z. Liu, L. L. Tang, J. F. Zong, Y. Huang, Y. Sun, Y. P. Mao, L. Z. Liu, A. H. Lin, and J. Ma, "Evaluation of sixth edition of AJCC staging system for nasopharyngeal carcinoma and proposed improvement," *Int J Radiat Oncol Biol Phys,* vol. 70, pp. 1115-23, 2008.

[13] A. Lin, H. M. Kim, J. E. Terrell, L. A. Dawson, J. A. Ship, and A. Eisbruch, "Quality of life after parotid-sparing IMRT for head-and-neck cancer: a prospective longitudinal study," *Int J Radiat Oncol Biol Phys,* vol. 57, pp. 61-70, 2003.
[14] F. M. Fang, W. L. Tsai, T. F. Lee, K. C. Liao, H. C. Chen, and H. C. Hsu, "Multivariate analysis of quality of life outcome for nasopharyngeal carcinoma patients after treatment," *Radiother Oncol,* vol. 97, pp. 263-9, 2010.
[15] I. Beetz, C. Schilstra, A. van der Schaaf, E. R. van den Heuvel, P. Doornaert, P. van Luijk, A. Vissink, B. F. van der Laan, C. R. Leemans, H. P. Bijl, M. E. Christianen, R. J. Steenbakkers, and J. A. Langendijk, "NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors," *Radiother Oncol,* vol. 105, pp. 101-6, 2012.
[16] A. Borque, G. Sanz, C. Allepuz, L. Plaza, P. Gil, and L. A. Rioja, "The use of neural networks and logistic regression analysis for predicting pathological stage in men undergoing radical prostatectomy: a population based study," *J Urol,* vol. 166, pp. 1672-8, 2001.
[17] J. H. Song, S. S. Venkatesh, E. A. Conant, P. H. Arger, and C. M. Sehgal, "Comparative analysis of logistic regression and artificial neural network for computer-aided diagnosis of breast masses," *Acad Radiol,* vol. 12, pp. 487-95, 2005.
[18] B. Eftekhar, K. Mohammad, H. E. Ardebili, M. Ghodsi, and E. Ketabchi, "Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data," *BMC Medical Informatics and Decision Making,* vol. 5, p. 3, 2005.