Breast Cancer Res Treat (2012) 135:221–229
DOI 10.1007/s10549-012-2174-6

CLINICAL TRIAL

Predicting two-year quality of life after breast cancer surgery using artificial neural network and linear regression models

Hon-Yi Shi · Jinn-Tsong Tsai · Yao-Mei Chen · Richard Culbertson · Hong-Tai Chang · Ming-Feng Hou

Received: 11 April 2012 / Accepted: 17 July 2012 / Published online: 27 July 2012

Abstract The purpose of this study was to validate the use of artificial neural network (ANN) models for predicting quality of life (QOL) after breast cancer surgery and to compare the predictive capability of ANNs with that of linear regression (LR) models. The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire and its supplementary breast cancer measure were completed by 402 breast cancer patients at baseline and at 2 years postoperatively. The accuracy of the system models were evaluated in terms of mean square error (MSE) and mean absolute percentage error (MAPE). A global sensitivity analysis was also performed to assess the relative significance of input parameters in the system model and to rank the variables in order of importance. Compared to the LR model, the ANN model generally had smaller MSE and MAPE values in both the training and testing datasets. Most ANN models had MAPE values ranging from 4.70 to 19.96 %, and most had high prediction accuracy. The ANN model also outperformed the LR model in terms of prediction accuracy. According to global sensitivity analysis, pre-operative functional status was the best predictor of QOL after surgery. Compared with the conventional LR model, the ANN model in the study was more accurate for predicting patient-reported QOL and had higher overall performance indices. Further refinements are expected to obtain sufficient performance improvements for its routine use in clinical practice as an adjunctive decision-making tool.

Keywords Breast cancer · Quality of life · Artificial neural network · Linear regression · Global sensitivity analysis

H.-Y. Shi
Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Kaohsiung 80708, Taiwan
e-mail: hshi@kmu.edu.tw

J.-T. Tsai
Department of Computer Science, National Pingtung University of Education, Pingtung 900, Taiwan

Y.-M. Chen
Faculty of Nursing, Kaohsiung Medical University, Kaohsiung 80708, Taiwan

Y.-M. Chen
Department of Nursing, Kaohsiung Municipal Hsiao-Kang Hospital, Kaohsiung 812, Taiwan

R. Culbertson
Department of Global Health Systems and Development, Tulane University, New Orleans, LA 70112, USA

H.-T. Chang
Division of General & Gastroenterological Surgery, Department of Surgery, Kaohsiung Veterans General Hospital, Kaohsiung 80708, Taiwan

M.-F. Hou
Division of General & Gastroenterological Surgery, Department of Surgery, Kaohsiung Medical University Hospital, Kaohsiung 80708, Taiwan

M.-F. Hou (✉)
Cancer Center, Kaohsiung Medical University Hospital and Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University, 100 Shi-Chuan 1st Road, Kaohsiung 80708, Taiwan, ROC
e-mail: mifeho@kmu.edu.tw

M.-F. Hou
National Sun Yat-Sen University-Kaohsiung Medical University Joint Research Center, 804 Kaohsiung, Taiwan

Abbreviations

BCS	Breast-conserving surgery
MRM	Modified radical mastectomy
TRAM	Transverse rectus abdominus muscle
QOL	Quality of life
ANNs	Artificial neural networks
LR	Linear regression
MLP	Multilayer perceptron
MSE	Mean square error
MAPE	Mean absolute percentage error
VSR	Variable sensitivity ratios

Introduction

Women with early stage breast carcinoma generally have three equally effective surgical options: breast-conserving surgery (BCS), modified radical mastectomy (MRM), or transverse rectus abdominus muscle (TRAM) flap surgery. Since the three procedures have comparable survival rates, patients typically select the procedure that optimizes quality of life (QOL) [1–3].

Artificial neural networks (ANNs) are complex and flexible nonlinear systems with properties not found in other modeling systems. These properties include robust performance in dealing with noisy or incomplete input patterns, high fault tolerance, and the capability to generalize from the input data [4–6]. The computational power of an ANN is derived from the distributed nature of its connections. Once a model is trained, it can be tested against novel records to predict outputs [4–6].

Although models proposed in the literature so far have contributed to the growing understanding of breast cancer surgery outcomes, they have had major shortcomings [7–10]. First, few studies of breast cancer outcomes have used longitudinal data for more than 2 years. Second, most studies have analyzed populations in the United States (US) or other countries, which may substantially differ from those in Taiwan. Third, no studies have considered group differences in factors other than outcomes such as age and nonsurgical treatment. Finally, almost all published articles agree that the essential issue of the internal validity (reproducibility) of the ANN and regression models has not been adequately addressed.

Therefore, the primary aim of the study was to validate the use of ANN models in predicting patient-reported QOL after breast cancer surgery, and the secondary aim was to compare the predictive capability of ANNs with that of linear regression (LR) models.

Materials and methods

Study design and population

The study included all patients who had been diagnosed and treated for incidental breast cancer between August, 2007 and September, 2009 at either of two participating tertiary academic hospitals in southern Taiwan. Patients who presented with curable diseases (i.e., no distant metastasis) were offered counseling regarding their surgical options (BCS, MRM, or TRAM flap surgery). After excluding patients with benign tumor ($n = 342$) or cognitive impairment ($n = 4$), 479 patients who gave written consent were enrolled in the study. At 2 years postoperatively, seventy-six patients were excluded due to loss to follow-up ($n = 57$) or refusal to participate ($n = 19$). The remaining 403 patients completed two surveys.

Instruments

The European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30 and QLQBR23 questionnaires were used to assess QOL [11, 12]. The Chinese versions of the EORTC QLQ-C30 and EORTC QLQ-BR23 have been validated in breast cancer surgical patients in Taiwan [13]. Since most symptom subscales of the QLQ-C30 and the QLQ-BR23 refer to systematic treatment effects, the analysis in this study was limited to the function subscales and global quality of life.

Before performing this study of human subjects, approval was obtained from all participating institutions. In all subjects, the QLQ-C30 and the QLQ-BR23 were administered by the same two research assistants before and after surgery.

System model development

The factors used in the LR model to predict long-term QOL of breast cancer surgery patients included both patient characteristics and hospital characteristics. The LR model can be formulated as the following linear equation:

$$\hat{Y} = \beta_0 + \beta_i X_i + \varepsilon_i, \quad i = 1, 2, \ldots, m.$$

where $\hat{Y}$ is the actual output value, β_0 is the intercept, β_i is the model coefficient parameter, X_i is the independent or input variable, ε_i is the random error, and m is the number of variables.

The ANN used in this study was a standard feed-forward, back-propagation neural network with three layers: an input layer, a hidden layer, and an output layer. The multilayer perceptron (MLP) network is an emerging tool for designing special classes of layered feed-forward networks [14]. The cross-validation approach typically used to optimize the time when an MLP network training session "stops" is to include one estimation subset for training the model and one validation subset for evaluating the model performance. A neural network is optimized using a training dataset. A separate test dataset is used to halt training to mitigate overfitting. The training cycle is repeated until the test error no longer decreases [6, 15].

Statistical analysis

The dataset was randomly divided into two sets: one set of 322 cases (80 % of the overall dataset) for training the model and another set of 81 cases for testing the model. The model was built using the training set. Patient characteristics and hospital characteristics were the independent variables, and the outcome (QOL) was the dependent variable. The LR and ANN models were then tested using the 81 cases in the testing dataset.

The model fit and prediction accuracy of the system models were measured in terms of mean square error (MSE) and mean absolute percentage error (MAPE), respectively. The prediction accuracy of a model is considered excellent if its MAPE value is lower than 10 %. Values between 10 and 20 %, between 20 and 50 %, and higher than 50 % are considered indicators of high, average, and low prediction accuracy [16]. The formulas used to calculate MSE and MAPE were

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2,$$

and

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|Y_i - \hat{Y}_i\right|}{Y_i} \times 100\,\%,$$

where n is the number of observations, Y_i is the desired (target) value of the ith observation, and $\hat{Y}_i$ is the actual output value of the ith observation.

The change rate was also used to compare model performance between the training and testing sets. This criterion was used to calculate the difference in MSE index between the test and the training sets so that the better model could be identified. Absolute value was defined as [(the MSE value from testing set—the MSE value from training set)/(the MSE value from training set)] × 100 %. Low change rates and low MSE values were considered indicators of good model performance.

The unit of analysis in this study was the individual breast cancer surgery patient. The data analysis was performed in several stages. First, continuous variables were tested for statistical significance by one-way analysis of variance (ANOVA), and categorical variables were tested by Fisher exact analysis. Univariate analyses were applied to identify significant predictors ($p < 0.05$). Second, STATISTICA 10.0 (StatSoft, Tulsa, OK) software was used to construct the MLP network model and the LR model of the relationship between the identified predictors and QOL. Finally, a global sensitivity analysis was also performed to assess the relative significance of input parameters in the system model and to rank the variables in order of importance. The global sensitivity of the input variables against the output variable was expressed as the ratio of the network error (variable sensitivity ratios, VSR) with a given input omitted to the network error with the input included. A ratio of one or lower indicates that the variable degrades network performance and should be removed [17].

Results

Table 1 shows the patient characteristics and hospital characteristics in this study. The mean age of the study population was 52.21 (±9.59) years. On average, 88.18 % of female patients were married, and the overall CCI was 0.59 ± 0.99. Of these 403 patients, eight patients in stage IV showed confirmed true metastatic disease, including two lung metastases, two liver metastases, and four bone metastases; and forty-eight patients showed a breast cancer history before receiving the surgical procedure. The significant variables ultimately selected for inclusion in the LR models were education, menopause status, surgical type, chemotherapy, radiotherapy, hormone therapy, post-operation LOS, complications, and pre-operation functional status ($p < 0.05$) (Table 2).

In this study of the MLP network, 80 % training and 20 % testing samples are randomly selected to analyze the database in each run. In order to make MLP learning perform better, the neuron activation functions for the hidden and output neurons availably are given as follows, such as identity, hyperbolic tangent, logistic sigmoid, exponential, and Sine. The optimal number of neurons in the hidden layer and the type of the activation functions are iteratively determined by developing 50 neural networks and observing the MSE index of the output error. The training process would continue training the network for as many cycles as needed so long as the training and testing errors are on the decrease, otherwise it would stop training as the test error increases. The ANN-based approaches provided the 3-layer networks and the relative weights of neurons used for predicting QOL. For example, the MLP 9-13-1 model for the QLQ-BR23 body image score prediction included nine inputs, one bias neuron in the input layer, 13 hidden neurons, one bias neuron in the hidden layer, and one output neuron (Table 3). The activation functions of logistic sigmoid and hyperbolic tangent were used in each neuron of the hidden layer and output layer, respectively.

For predicting QOL, the ANN model had relatively larger change rates and MSE values in the testing set with the exception of MSE for the testing set at year two (Tables 4, 5). Apparently, the ANN model also outperformed the LR model in terms of predictive accuracy. Most MAPE values obtained by the ANN model were lower than

Table 1 Patient and hospital characteristics of the study ($N = 403$)

Variables	Mean ± SD	*N* (%)
Patient characteristics		
Age at operation (years)	52.21 ± 9.59	
Married		
No		48 (11.82)
Yes		355 (88.18)
Education (years)	9.43 ± 4.58	
Living with immediate family		
No		16 (3.97)
Yes		387 (96.03)
Body mass index (kg/m^2)	23.85 ± 3.61	
Smoker		
No		389 (96.53)
Yes		14 (3.47)
Drinker		
No		391 (97.02)
Yes		12 (2.98)
Menopause status		
No		200 (49.63)
Yes		203 (50.37)
Number of fetuses (cases)	2.35 ± 1.17	
Breast cancer history		
No		355 (88.09)
Yes		48 (11.91)
Other breast disease history		
No		333 (82.63)
Yes		70 (17.37)
Charlson co-morbidity index	0.59 ± 0.99	
Tumor pathology differentiation		
High		56 (13.90)
Medium		270 (67.00)
Low		77 (19.10)
Tumor stage		
Stage 0/I		159 (39.45)
Stage II		147 (36.48)
Stage III/IV		97 (24.17)
Hospital characteristics		
Surgical procedure		
MRM		234 (58.06)
BCS		113 (28.04)
TRAM		56 (13.90)
Operation time (min)	166.92 ± 107.36	
Anesthesia time (min)	197.71 ± 117.06	
ASA class		
I		40 (9.93)
II		312 (77.42)
III		51 (12.65)
Chemotherapy		
No		103 (25.56)
Yes		300 (74.44)
Radiotherapy		
No		232 (57.57)
Yes		171 (42.43)
Hormone therapy		
No		228 (56.57)
Yes		175 (43.43)
Post-operation LOS (days)	3.09 ± 1.49	
Post-hospitalization 30 days		
No		324 (80.40)
Yes		79 (19.60)
Complications		
No		351 (87.10)
Yes		52 (12.90)

MRM modified radical mastectomy; *BCS* breast-conserving surgery; *TRAM* transverse rectus abdominus muscle mastectomy with reconstruction; *ASA* American society of anesthesiologists; *LOS* length of stay; *SD* standard deviation

20 %, which indicated that the ANN model had excellent accuracy for predicting QOL.

Table 6 presents the VSR values for the outcome variable (QOL) in relation to the three most influential variables. In the global sensitivity analysis, the most influential (sensitive) parameter in terms of its effects on most QLQ-BR23 and QLQ-C30 subscales was pre-operative functional status followed by surgical type. All VSR values exceeded one, indicating that the network performed better when all variables were considered.

In order to verify the predictive accuracy of the models, the 40 datasets shown in Table 7 were collected. Compared to the LR model, the ANN model consistently obtained higher performance indices in the QLQ-BR23 and QLQ-C30 subscales.

Discussion

Floyd et al. [18] was the first to develop an ANN model for predicting breast cancer based on mammographic findings. They concluded that the ANN model could be trained to predict malignancy based on mammographic findings with accuracy exceeding that of experienced radiologists. Ayer et al. [19] constructed LR and ANN models for estimating breast cancer risk based on mammographic descriptors and demographic risk factors. They concluded that the ANN model can be viewed as a generalization of the LR model and that the main advantage of ANN models over LR models is their hidden layers of nodes. Orr developed a

Table 2 Coefficients of selected significant variables in each quality of life subscale of linear regression model ($N = 403$)[a]

Variable	BRBI	BRSEF	BRSEE	BRFU	QL	PF	RF	EF	CF	SF
Education	0.07	0.04	0.19	−0.02	0.14	0.13	8.31	0.11	0.27	0.07
Menopause status										
Yes versus no	2.02	0.57	2.39	−0.22	0.38	−0.17	−107.67	0.24	−0.31	0.58
Surgical type										
BCS versus MRM	−1.38	−2.27	−0.48	0.29	1.76	−1.17	1.50	0.41	0.16	−1.24
TRAM versus MRM	1.25	−0.20	−0.70	0.03	1.69	−1.69	−8.87	1.47	1.40	−1.05
Chemotherapy										
Yes versus no	1.93	0.44	5.97	−1.07	0.22	0.31	−55.09	−0.86	−0.10	0.76
Radiotherapy										
Yes versus no	4.66	0.28	−3.68	0.51	0.62	−0.06	99.03	1.19	1.67	1.49
Hormone therapy										
Yes versus no	−1.37	−0.11	−2.99	−0.59	0.88	0.89	−10.13	−0.42	−0.13	−0.07
Post-operation LOS	−0.09	0.79	1.27	−0.34	−0.15	0.19	23.26	−0.46	−0.75	−0.89
Complications										
Yes versus no	−21.41	−1.09	−0.55	0.15	−0.84	−0.77	−119.57	0.74	0.06	−0.84
Pre-operation functional status	−0.27	−0.25	−0.95	−0.04	−0.11	−0.24	−9.17	−0.15	−0.24	−0.22
Constant	16.31	14.71	56.98	2.61	4.01	18.49	664.70	11.27	16.76	17.54

BRBI QLQ-BR23 body image; *BRSEF* QLQ-BR23 sexual functioning; *BRSEE* QLQ-BR23 sexual enjoyment; *BRFU* QLQ-BR23 future perspective; *QL* QLQ-C30 global quality of life; *PF* QLQ-C30 physical functioning; *RF* QLQ-C30 role functioning; *EF* QLQ-C30 emotional functioning; *CF* QLQ-C30 cognitive functioning; *SF* QLQ-C30 social functioning; *MRM* modified radical mastectomy; *BCS* breast-conserving surgery; *TRAM* transverse rectus abdominus muscle mastectomy with reconstruction

[a] All regression coefficients are statistically significant ($p < 0.05$)

Table 3 Artificial neural network models at different subscales

Subscale	Net[a]
QLQ-BR23 body image	9–13–1
QLQ-BR23 sexual functioning	9–9–1
QLQ-BR23 sexual enjoyment	9–5–1
QLQ-BR23 future perspective	9–18–1
QLQ-C30 global quality of life	9–11–1
QLQ-C30 physical functioning	9–4–1
QLQ-C30 role functioning	9–13–1
QLQ-C30 emotional functioning	9–18–1
QLQ-C30 cognitive functioning	9–10–1
QLQ-C30 social functioning	9–4–1

[a] Input layer–hidden layer–output layer

simplified and standardized method of classifying patients with abnormal mammograms by incorporating quantitative risk assessment [20]. His performance comparisons of ANN models and conventional LR models used for mammographic classification showed better discrimination in the ANN model. From a practical standpoint, however, the models showed similar performance in identifying malignant cases misclassified by clinical impression.

This study confirmed that, compared to LR models, ANN models are dramatically more accurate in predicting patient-reported outcomes (QOL). To the best of our knowledge, this study is the first to use ANNs for analyzing predictors of QOL after breast cancer surgery. This model was tested against actual outcomes obtained by models constructed using identical inputs, including a neural network model and a linear regression model. We also showed that, given the same numbers of inputs for patient characteristics and hospital characteristics and the same two outcome measures, the predictive accuracy of ANN is superior to that of LR.

Multiple outcome-predicting models have been developed with conventional statistical procedures, but their application at the individual level is hampered by the highly interdependent clinical variables involved, which may potentially interact with each other and have reciprocal enhancing effects [6, 10]. Hence, conventional statistical approaches have intrinsic limitations in handling this complex nonlinear information [4–6].

The ANNs are adaptive models that use a dynamic approach for analyzing outcome risks and can modify the internal structure in relation to a functional objective [4–6]. Although conventional statistics reveal significant parameters only for the overall population, ANNs include parameters that are significant at the individual patient level even if they are not significant in the overall population [5, 6]. We believe that the large and homogeneous

Table 4 Comparison of artificial neural network (ANN) model and linear regression (LR) model in predicting QLQ-BR23 subscale scores

Index	Model	Training set (A)	Testing set (B)	Change rate[a]
QLQ-BR23 body image score				
MSE	ANN	66.67	84.85	27.27 %
	LR	70.00	88.89	26.99 %
MAPE	ANN	17.14 %	19.57 %	–
	LR	20.46 %	28.23 %	–
QLQ-BR23 sexual functioning score				
MSE	ANN	66.67	50.00	25.00 %
	LR	72.22	57.14	15.08 %
MAPE	ANN	12.50 %	8.79 %	–
	LR	22.24 %	10.32 %	–
QLQ-BR23 sexual enjoyment score				
MSE	ANN	75.00	50.00	33.33 %
	LR	83.33	66.67	19.99 %
MAPE	ANN	16.81 %	8.84 %	–
	LR	27.31 %	12.83 %	–
QLQ-BR23 future perspective score				
MSE	ANN	83.33	68.18	18.18 %
	LR	90.32	75.00	16.96 %
MAPE	ANN	19.96 %	16.10 %	–
	LR	34.71 %	22.87 %	–

MSE mean square error; *MAPE* mean absolute percentage error

[a] Change rate = $|[(B - A)/(A)]| \times 100\,\%$

Table 5 Comparison of artificial neural network (ANN) model and linear regression (LR) model in predicting QLQ-C30 subscale scores

Index	Model	Training set (A)	Testing set (B)	Change rate[a]
QLQ-C30 global quality of life score				
MSE	ANN	29.31	14.81	49.47 %
	LR	34.67	18.24	47.38 %
MAPE	ANN	10.75 %	6.14 %	–
	LR	14.63 %	8.06 %	–
QLQ-C30 physical functioning score				
MSE	ANN	83.87	66.67	20.51 %
	LR	92.77	77.59	16.36 %
MAPE	ANN	17.14 %	15.81 %	–
	LR	35.57 %	19.31 %	–
QLQ-C30 role functioning score				
MSE	ANN	18.68	6.24	66.60 %
	LR	24.07	11.20	53.47 %
MAPE	ANN	7.09 %	4.84 %	–
	LR	9.31 %	6.50 %	–
QLQ-C30 emotional functioning score				
MSE	ANN	83.87	75.00	4.54 %
	LR	92.77	88.89	4.18 %
MAPE	ANN	17.43 %	16.57 %	–
	LR	40.56 %	24.23 %	–
QLQ-C30 cognitive functioning score				
MSE	ANN	75.00	50.00	33.33 %
	LR	83.33	59.09	29.09 %
MAPE	ANN	16.01 %	10.46 %	–
	LR	18.14 %	11.19 %	–
QLQ-C30 social functioning score				
MSE	ANN	17.47	5.84	66.57 %
	LR	17.47	10.42	16.96 %
MAPE	ANN	8.64 %	4.70 %	–
	LR	9.04 %	5.64 %	–

MSE mean square error; *MAPE* mean absolute percentage error

[a] Change rate = $|[(B - A)/(A)]| \times 100\,\%$

dataset in the present study, which included all demographic and clinical variables shown to affect QOL in previous linear regression models, provided a sufficiently robust basis for training the network [5, 6].

Throughout this two-year follow-up study, the best single predictor of QOL subscale scores was pre-operation functional status, which is consistent with reports that pre-operation functional scores are the best predictors of postoperative QOL [2, 21]. Therefore, effective counseling is essential for apprising patients of expected post-surgery impairments. If QOL outcomes are considered as benchmarks then pre-operation functional status, which is a major predictor of postoperative QOL, is crucial. Patients should also be advised that their postoperative QOL might depend not only on the success of their operations, but also on their pre-operation functional status.

Furthermore, recent findings suggest that BCS outperforms MRM for measuring role functioning, emotional functioning, cognitive functioning, and body image [2]. Compared with the BCS groups, however, the TRAM groups revealed significantly larger subjective improvements in physical functioning, emotional functioning, sexual functioning, and sexual enjoyment. One study found that aspects of QOL, other than body image, were no better in women who underwent breast-conserving surgery or mastectomy with reconstruction than in women who had mastectomy alone [22]. Mastectomy with reconstruction was associated with greater mood disturbance and poorer health. However, the results of a 5-year prospective study on QOL following breast-conserving surgery or mastectomy indicated that mastectomy patients had a significantly worse body image, role, and sexual functioning, and their lives were more disrupted [23].

In addition, to reduce the risk of recurrence and death, breast cancer patients usually receive systemic therapies (chemotherapy, hormone therapy, radiotherapy, and biological treatments) after surgery. Several studies

Table 6 Global sensitivity analysis of QLQ-BR23 and QLQ-C30 subscales of artificial neural network (ANN) model

Subscale	First (VSR)	Second (VSR)	Third (VSR)
QLQ-BR23 body image	Surgical type (1.54)	Pre-operation functional status (1.44)	Radiotherapy (1.35)
QLQ-BR23 Sexual functioning	Pre-operation functional status (1.86)	Surgical type (1.56)	Complication (1.21)
QLQ-BR23 Sexual enjoyment	Pre-operation functional status (6.49)	Complication (1.27)	Surgical type (1.17)
QLQ-BR23 Future perspective	Surgical type (4.20)	Chemotherapy (3.45)	Pre-operation functional status (2.37)
QLQ-C30 Global quality of life	Pre-operation functional status (2.21)	Surgical type (2.19)	Complication (1.49)
QLQ-C30 Physical functioning	Pre-operation functional status (2.60)	Complication (1.79)	Surgical type (1.71)
QLQ-C30 Role functioning	Pre-operation functional status (82.89)	Complication (12.54)	Chemotherapy (9.46)
QLQ-C30 Emotional functioning	Pre-operation functional status (1.36)	Surgical type (1.14)	Complication (1.10)
QLQ-C30 Cognitive functioning	Pre-operation functional status (1.86)	Complication (1.68)	Surgical type (1.55)
QLQ-C30 Social functioning	Surgical type (8.97)	Complication (7.91)	Pre-operation functional status (6.86)

VSR variable sensitivity ratios

Table 7 Comparison of performance indices of artificial neural network (ANN) model and linear regression (LR) model for predicting QLQ-BR23 and QLQ-C30 subscale scores based on forty new datasets

Subscale	ANN model						LR model					
	Sensitivity	1-Specificity	PPV	NPV	Accuracy	AUC	Sensitivity	1-Specificity	PPV	NPV	Accuracy	AUC
QLQ-BR23												
BRBI	100.00	100.00	1.00	1.00	100.00	1.00	92.86	92.86	0.92	0.92	93.33	0.92
BRSEF	95.83	90.91	0.96	0.91	90.00	0.92	75.00	90.91	0.75	0.91	86.67	0.83
BRSEE	100.00	95.83	0.86	1.00	96.67	0.98	100.00	95.83	0.86	1.00	96.67	0.98
BRFU	66.67	100.00	1.00	0.92	93.33	0.83	33.33	91.67	0.50	0.85	80.00	0.63
QLQ-C30												
QL	100.00	81.25	0.82	1.00	90.00	0.91	92.86	81.25	0.81	0.93	86.67	0.87
PF	66.67	95.83	0.80	0.92	90.00	0.81	66.67	95.83	0.80	0.92	90.00	0.81
RF	100.00	100.00	1.00	1.00	100.00	1.00	100.00	95.83	0.86	1.00	96.67	0.98
EF	66.67	95.83	0.86	0.89	86.67	0.87	66.67	94.44	0.89	0.81	83.33	0.81
CF	66.67	100.00	1.00	0.92	93.33	0.92	40.00	96.00	0.67	0.89	86.67	0.68
SF	100.00	100.00	1.00	1.00	100.00	1.00	83.33	100.00	1.00	0.96	96.67	0.92

BRBI QLQ-BR23 body image; *BRSEF* QLQ-BR23 sexual functioning; *BRSEE* QLQ-BR23 sexual enjoyment; *BRFU* QLQ-BR23 future perspective; *QL* QLQ-C30 global quality of life; *PF* QLQ-C30 physical functioning; *RF* QLQ-C30 role functioning; *EF* QLQ-C30 emotional functioning; *CF* QLQ-C30 cognitive functioning; *SF* QLQ-C30 social functioning; *PPV* positive predictive value; *NPV* negative predictive value; *AUC* area under the curve

evaluated QOL on breast cancer patients who have received systemic therapies [1–3]. Chemotherapy has considerable effects on QOL for breast cancer patients. Notedly, a complication is a well-recognized risk factor with adverse outcomes in breast cancer surgery. Our statistical data also show a strong and positive association with poor QOL, which is consistent with previous findings [2, 24].

Although all research questions were satisfactorily addressed, several limitations are noted. First, this study collected data for breast cancer surgery patients who had been under the supervision of two surgeons in two different medical centers, each of whom had performed the highest volume of breast cancer surgery procedures in his respective hospital during the previous 20–30 years. This sample selection procedure ensured that patient outcome data would not be affected by surgeons with limited experience. By focusing the analysis on procedures performed by these two surgeons, the results of this study are more representative of all breast cancer patients compared to one analyzing those performed by a single surgeon. However, a notable limitation is that the first patient in the prospective patient cohort was enrolled in 2007. Therefore, depending on their inclusion date, some surveyed patients had a longer follow-up than others did, which may have caused selection bias. Nonetheless, in most QOL subscales, the characteristics of subjects who continuously participated throughout this 2-year study did not significantly differ from those of subjects who died or dropped out during the study (data not shown).

Conclusions

Compared with the conventional multivariate LR model, the ANN model in the study was more accurate in predicting patient-reported QOL and had higher overall performance indices. The global sensitivity analysis also showed that preoperation functional status is the most important predictor of the QLQ-BR23 and the QLQ-C30 after breast cancer surgery. The predictors analyzed in this study could be addressed in pre-operative and postoperative health care consultations to educate candidates for breast cancer surgery in the expected course of recovery and expected functional outcomes. Further studies of this model may consider the effect of a more detailed database that includes complications and clinical examination findings as well as more detailed outcome data. Hopefully, the model will evolve into an effective adjunctive clinical decision-making tool.

Acknowledgments This work was supported by the National Science Council, Taiwan, Republic of China, under grant numbers NSC99-2314-B-037-069-MY3 and by the Department of Health, Executive Yuan, under Grant Numbers DOH101-TD-C-111-002.

Conflicts of interest The authors have no personal or professional conflicts of interest in the publication of this study.

References

1. Dawood S, Hu R, Homes MD, Collins LC, Schnitt SJ, Connolly J, Colditz GA, Tamimi RM (2011) Defining breast cancer prognosis based on molecular phenotypes: results from a large cohort study. Breast Cancer Res Treat 126:185–192
2. Shi HY, Uen YH, Yen LC, Culbertson R, Juan CH, Hou MF (2011) Two-year quality of life after breast cancer surgery: a comparison of three surgical procedures. Eur J Surg Oncol 37:695–702
3. van den Hurk CJ, Eckel R, van de Poll-Franse LV, Coebergh JW, Nortier JW, Hölzel D, Breed WP, Engel J (2011) Unfavourable pattern of metastases in M0 breast cancer patients during 1978–2008: a population-based analysis of the Munich Cancer Registry. Breast Cancer Res Treat 128:795–805
4. Tu JV (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 49:1225–1231
5. Zou J, Han Y, So SS (2008) Overview of artificial neural networks. Methods Mol Biol 458:15–23
6. Sandberg IW, Lo JT, Fancourt CL, Principe JC, Katagiri S, Haykin S (2001) Nonlinear dynamical systems: feedforward neural network perspectives. Wiley, New York
7. Giordano A, Giuliano M, De Laurentiis M, Eleuteri A, Iorio F, Tagliaferri R, Hortobagyi GN, Pusztai L, De Placido S, Hess K, Cristofanilli M, Reuben JM (2011) Artificial neural network analysis of circulating tumor cells in metastatic breast cancer patients. Breast Cancer Res Treat 129:451–458
8. Lancashire LJ, Powe DG, Reis-Filho JS, Rakha E, Lemetre C, Weigelt B, Abdel-Fatah TM, Green AR, Mukta R, Blamey R, Paish EC, Rees RC, Ellis IO, Ball GR (2010) A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. Breast Cancer Res Treat 120: 83–93
9. Foukakis T, Fornander T, Lekberg T, Hellborg H, Adolfsson J, Bergh J (2011) Age-specific trends of survival in metastatic breast cancer: 26 years longitudinal data from a population-based cancer registry in Stockholm, Sweden. Breast Cancer Res Treat 130:553–560
10. Zujewski JA, Harlan LC, Morrell DM, Stevens JL (2011) Ductal carcinoma in situ: trends in treatment over time in the US. Breast Cancer Res Treat 127:251–257
11. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A (2006) EORTC QLQ C30 scoring manual, 3rd edn. EORTC, Brussels
12. Sprangers MA, Groenvold M, Arraras JI, Franklin J, te Velde A, Muller M, Franzini L, Williams A, de Haes HC, Hopwood P, Cull A, Aaronson NK (1996) The European Organization for Research and Treatment of Cancer breast cancer-specific quality-of-life questionnaire module: first results from a three-country field study. J Clin Oncol 14:2756–2768
13. Chie WC, Chang KJ, Huang CS, Kuo WH (2003) Quality of life of breast cancer patients in Taiwan: validation of the Taiwan Chinese version of the EORTC QLQ-C30 and EORTC QLQ-BR23. Psychooncology 12:729–735
14. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. In: Rumelhart DE, McCleland JL (eds) Parallel distributed processing: explorations in the microstructure of cognition. MIT Press, Cambridge, pp 318–362
15. Haykin S (1999) Neural networks: a comprehensive foundation, 2nd edn. Prentice-Hall, Englewood Cliffs
16. Woods LM, Coleman MP, Lawrence G, Rashbass J, Berrino F, Rachet B (2011) Evidence against the proposition that "UK cancer survival statistics are misleading": simulation study with national cancer registry data. Br Med J 342:d3399
17. Hunter A, Kennedy L, Henry J, Ferguson I (2000) Application of neural networks and sensitivity analysis to improved prediction of trauma survival. Comput Methods Programs Biomed 62:11–19

18. Floyd CE Jr, Lo JY, Yun AJ, Sullivan DC, Kornguth PJ (1994) Prediction of breast cancer malignancy using an artificial neural network. Cancer 74:2944–2948
19. Ayer T, Chhatwal J, Alagoz O, Kahn CE Jr, Woods RW, Burnside ES (2010) Informatics in radiology: comparison of logistic regression and artificial neural network models in breast cancer risk estimation. Radiographics 30:13–22
20. Orr RK (2001) Use of an artificial neural network to quantitate risk of malignancy for abnormal mammograms. Surgery 129: 459–466
21. Rottmann N, Dalton SO, Christensen J, Frederiksen K, Johansen C (2010) Self-efficacy, adjustment style and well-being in breast cancer patients: a longitudinal study. Qual Life Res 19:827–836
22. Nissen MJ, Swenson KK, Ritz LJ, Farrell JB, Sladek ML, Lally RM (2001) Quality of life after breast carcinoma surgery: a comparison of three surgical procedures. Cancer 91:1238–1246
23. Engel J, Kerr J, Schlesinger-Raab A, Sauer H, Halzel D (2004) Quality of life following breast-conserving therapy or mastectomy: results of a 5-year prospective study. Breast J 10:223–231
24. Deshpande AD, Sefko JA, Jeffe DB, Schootman M (2011) The association between chronic disease burden and quality of life among breast cancer survivors in Missouri. Breast Cancer Res Treat 129:877–886