ESANN'2005 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 27-29 April 2005, d-side publi., ISBN 2-930307-05-6.

Artificial Intelligence Techniques for the Prediction of Bladder Cancer Progression

M.F. Abbod[1] and J.W.F. Catto[2], M. Chen[1], D.A. Linkens[1] and F.C. Hamdy[2]
[1]Department of Automatic Control and Systems Engineering and [2]Academic Urology Unit, University of Sheffield, United Kingdom

Abstract. New techniques for the prediction of tumour behaviour are needed since statistical analysis has a poor accuracy and is not applicable to the individual. Artificial Intelligence (AI) may provide these suitable methods. We have compared the predictive accuracies of neuro-fuzzy modelling (NFM), artificial neural networks (ANN) and traditional statistical methods, for the behaviour of bladder cancer. Experimental molecular biomarkers, including p53 expression and gene methylation, and conventional clinicopathological data were studied in a cohort of 122 patients with bladder cancer. For all 3 methods, models were produced to predict the presence and timing of a tumour progression. Both methods of AI predicted progression with an accuracy ranging from 88-100%. This was superior to logistic regression. NFM appeared better than ANN at predicting the timing of progression.

1 Introduction

Transitional cell carcinoma (TCC) of the bladder is the 4th commonest cancer amongst men in the UK [1]. At presentation, 70% of TCC are superficial and non-invasive, which can be managed by local endoscopic resection and intra-vesical chemotherapy. Following treatment, these tumours require cystoscopic surveillance [2] as 50% will recur as similar non-invasive lesions, and a smaller percentage (20%) will progress to muscle invasion. Muscle invasive tumours have a poor prognosis (50% 5 year survival rates) and require radical therapy if cure is to be achieved [3]. Following radical treatment, adjuvant chemotherapy can be used to reduce relapse and possibly mortality rates in selected patients with the highest chance of relapse [4].

The accurate prediction of future cancer behaviour would be of obvious benefit to both the patient and the physician. Patients with non-relapsing tumours could be safely reassured and discharged, whilst relapsing tumours could be treated more aggressively. The most reliable predictors of tumour behaviour are the pathological stage and grade at diagnosis (TNM classification) [5]. Specific tumours also have additional prognostic information; including lymph node status in invasive disease and recurrence in superficial disease. Whilst these parameters stratify patients into subgroups, it is impossible to predict individual tumour behaviour. The development of molecular medicine has yielded many new molecules that may be useful as predictive biomarkers. Some of the most biologically promising markers are the p53 and gene methylation. The p53 gene is mutated in over 50% of human cancers [7] and has been shown to predict recurrence and survival in bladder cancer [8]. Gene

methylation occurs in the majority of tumours and is associated with a poor outcome [9].

A solution to the problem of predicting tumour behaviour lies potentially within the interpretation of data. Traditional statistical methods, e.g. logistic regression, produce probabilities of behaviour, which may be applicable to a population but not predictive for an individual. Furthermore, their predictions are only accurate in 70% of tumours using the TNM classification [7]. By using Artificial Intelligence (AI) methodologies, such as artificial neural networks (ANN) and neuro-fuzzy modelling (NFM), complex relationships between dependant and independent variables, in a population whose distribution may not be normal, can be identified. As a result prediction of biological behaviour from both physiological and pathological data can be performed.

ANN, of which the most commonly used is the Multi-Layer Perceptron, have been applied to clinical medicine since 1989 [8]. They consist of three layers; an input layer for data entry, a hidden layer of networked neurones and an output layer [10]. Previous authors have shown that ANN are superior to standard statistical analysis in the diagnosis of chest pain [11], the TNM staging system at predicting breast and colorectal cancer outcomes [7] and predicting progression of poorly differentiated superficial bladder tumours. However, ANN are not without problems. They can be 'over-trained' to learn the inherent variation ('noise') of a sample population and the network is hidden within a functional 'black box'. Thus, it is difficult to gain insight into the solution used to resolve the clinical data; making subsequent analysis (to ensure clinical sense prevails) and interrogation of new variables almost impossible. NFM is an alternate AI method, without many of these drawbacks of ANN.

2 Artificial Intelligence Modelling

Two models were developed for both ANN and NFM. A *Classifier* predicted the likelihood of a tumour relapse (yes or no), before a second model, used as a *Predictor*, predicted the timing of this relapse (months after surgery). These two models were combined together in series; thus predicting if and when a relapse would occur. Those tumours without relapse were therefore excluded from the *Predictor* model. To discover the value of the putative molecular biomarkers, the data were analysed 4 ways for each model. For the first analysis (A), only the standard clinicopathological data were studied (cancer stage, grade, age, sex, smoking, other cancers). A second analysis (B) was then performed, which included the additional p53 molecular putative biomarker. The third analysis (C) was based on the standard clinicopathological data plus the degree of methylation in a tumour (as a percentage), while the last analysis (D) was the same as (C) plus methylation at the retinoic acid receptor – β gene (RARB; a particularly sensitive predictor of tumour behaviour).

For both of the AI methods, the input variables used to train and test the models are shown in Table 1. The output variable from each model was either the presence (*Classifier*) or the timing of a tumour relapse (*Predictor*). Both AI methods were programmed using commercially available software, Matlab. There were 15 hidden neurones within the ANN.

The patients were randomly divided into 10 aliquots (each of 6-7 patients). For each model, the ANN was trained on 90% of these batches, before testing on the final 10%. This was then repeated until all batches had been used to test the model, so called 'ensembling' [12]. For each session, 10 models were generated and the best selected. Using this 'best fit' model, the data were then re-tested to produce the final results. To minimise over-training, an initial session with a validation step was performed. The NFM analyses were performed on commercially available software, Matlab. The fuzzy-logic predictions were performed on in-house software [13]. The patients were analysed on the same data selection method performed for the ANN methods. To obtain a probability of tumour relapse using traditional statistical analysis, logistic regression (LR) was also performed. Previous authors have concluded this to be the traditional statistical method, with which ANN and NFM should be compared [14].

	Input Variables	Scoring		
Analysis A	Stage	Ta	T1	T2-4
	Grade	1	2	3
	Age	In years		
	Sex			
	Smoking Exposure	Pack years		
	Previous Cancers (non-TCC)	0 = none	1 = 1	2 = >2
B, C &D	p53	0 = normal	1 = abnormal	
	methylnation	percentage		
	RARB	0 = abnormal	1 = normal	

Table 1. Input variables for the modelling methods. Four analyses were performed, (A-D). For analysis (A) the inputs were the 6 conventional clinicopathological data. For analysis (B) there was a p53 additional input, analysis (C) includes the conventional inputs plus %met, and analysis (D) is the same as (C) plus RARB.

3 Modelling Results

The 122 patients with TCC investigated represented a typical UK population (median age = 70 years, 65% male and 62% smokers). Smoking was significantly related to more advanced disease rather than non-smoking.

The results of the *Classifier* models generated using ANN and NFM are shown in Table 2. The accuracies for ANN and NFM are 93% and 98% for analysis (B) and 90% and 100% for analysis (D), respectively. The results for the *Predictor* models are shown, in comparison with LR, in Table 3. In each case the difference between the actual and predicted time of progression is shown as a root mean square value (RMS). In all categories the AI models perform better than LR, and NFM is more accurate than ANN. As can be seen, ANN and NFM are significantly superior to LR. When ANN and NFM are directly compared, NFM is significantly better than ANN at predicting tumour progression. The predictions of all 3 methods are shown graphically, as scatter plots in Figure 1.

		Sensitivity	Specificity	Accuracy		Sensitivity	Specificity	Accuracy
A	ANN	81	95	89	NFM	88	99	94
B		87	97	93		100	97	98
C		83	100	92		100	100	100
D		80	100	90		100	100	100

Table 2. Risk of tumour progression: results for the AI classifier models. The table shows the results of the classifiers' prediction of a tumour relapse (yes or no). In analysis (B, C & D), the additional putative molecular biomarkers are included, altering the performance of each model.

	ANN				NFM				LR
case	training	validation	testing	total	training	validation	testing	total	total
A	10.05	15.71	25.95	9.01	5.08	19.63	14.12	5.31	13.42
B	9.44	12.02	21.54	9.85	1.42	26.53	25.37	5.18	13.26
C	9.79	13.93	23.71	8.01	2.29	31.99	25.37	5.17	19.12
D	9.08	13.33	18.06	8.35	3.26	17.81	21.77	5.09	18.12

Table 3. Time to tumour progression: results of the AI Predictor models and logistic regression. The results of the AI models predictions and logistic regression probabilities of time to tumour progression are shown, as root mean squares (difference between predicted and actual time of relapse). For ANN and NFM, after training and testing, the data was then analysed for the overall best-fit model.

4 Discussion

The accurate prediction of an individual patient's tumour response to treatment is a Holy Grail of oncology. Here we have shown that NFM can predict tumour behaviour with greater accuracy than both ANN and LR. Until the advent of AI, the best method of predicting tumour behaviour was logistic regression. Using Figure 1, if LR were applied in clinical practice, those patients with late relapsing tumours (over 40 months) would have had their most intensive cystoscopic surveillance too early for their actual relapse.

Previous workers have shown ANN can predict tumour behaviour more accurately than LR [7, 15-18] and clinicians. We have again confirmed that ANN provides a powerful and accurate predictive method. However, unlike these previous studies, we have been able to compare ANN with NFM. NFM is a relatively novel modelling technique, with few previous reports in the field of cancer prediction [19]. Both this previous report and our current study have shown that NFM produces a significantly more accurate prediction than ANN and LR (Table 3). In addition to accuracy, NFM has other benefits over ANN. Unlike the 'black-box' phenomena of ANN, which makes interpretation of the rules taken to solve a problem impossible, the NFM approach is transparent (Figure 2). By 'defuzzification', qualitative modelling figures can be translated into understandable medical terms. Thus NFM can incorporate expertise and allow predictions of outcome that result from changes in the value of individual inputs. These features make NFM an important tool, with extensive clinical applications. Whilst our predictions have been modelled using bladder cancer, these methods are transferable to many other human malignancies.

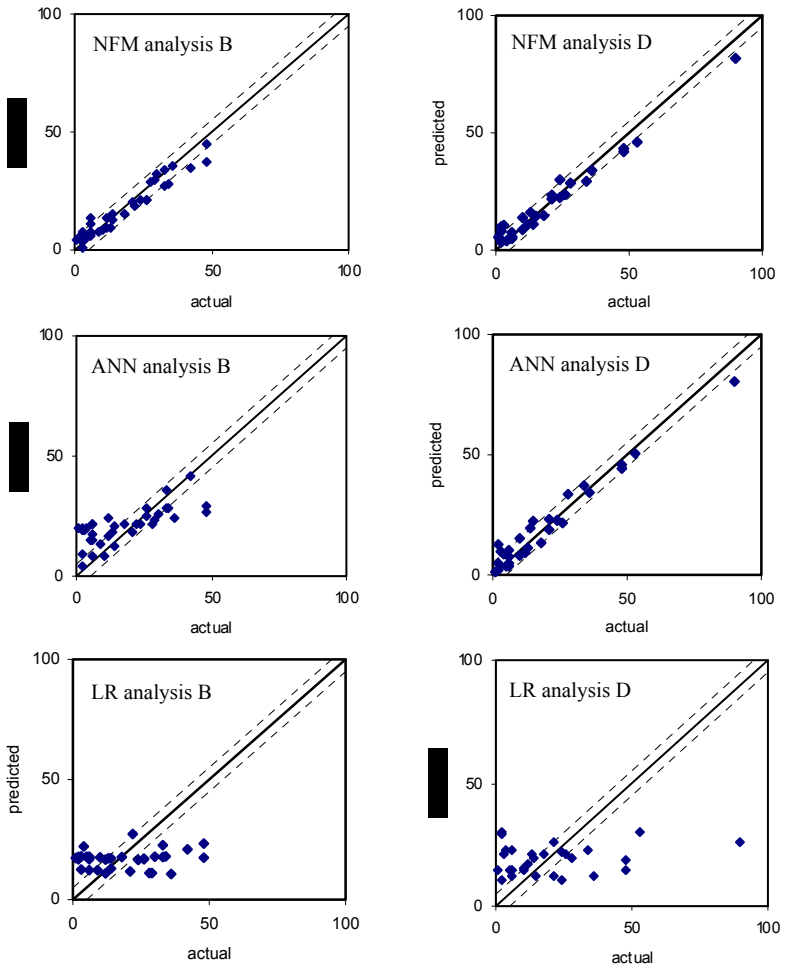

Figure 1. Scatter plots of actual and predicted progression times. In all six graphs each point represents the actual time of tumour progression (X axis; 0-100 months after surgery) against the predicted time of tumour progression (Y axis; 0-100 months after surgery). Continuous line indicates the accuracy of each plot. Dashed lines indicate the ±10% range.

References

[1] M. Quinn, Cancer trends in England and Wales 1950-1999. SMPs No. 66 ed: The Stationary Office, 2000.

[2] R.R. Hall, M.K. Parmar, A.B. Richards and P.H. Smith. Proposal for changes in cystoscopic follow up of patients with bladder cancer and adjuvant intravesical chemotherapy. *British Medical Journal,* 308(6923):257-260, 1994.

[3] J.P. Stein, G. Lieskovsky, R.J. Cote, et al., Radical cystectomy in the treatment of invasive bladder cancer; long term results in 1054 patients. *Journal of Clinical Oncology,* 19:666-675, 2001.

[4] D. Raghaven, D. Quinn, D.G. Skinner and J.P. Stein, Surgery and adjunctive chemotherapy for invasive bladder cancer. *Surgical Oncology,* 11:55-63, 2002.

[5] L.H. Sobin, C.H. Wittekind, TNM classification of malignant tumours. 5th edition New York (NY): John Wiley and sons, 1997.

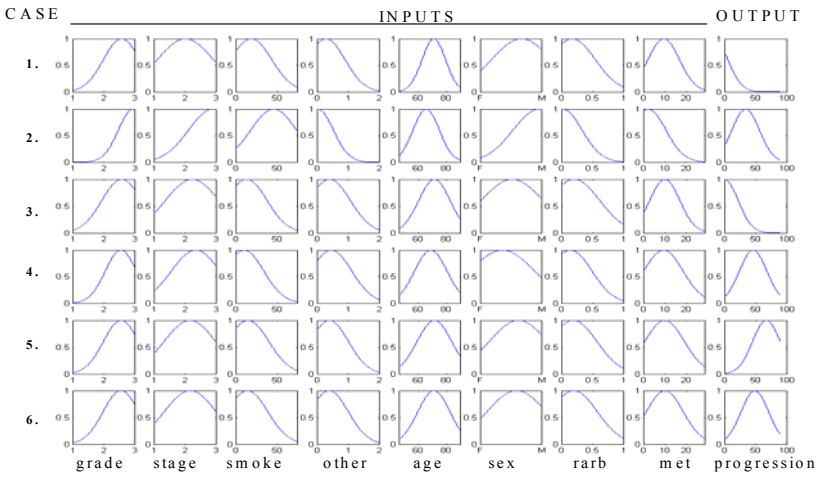

Figure 2. Neuro-Fuzzy Modelling output, after defuzzification. The above diagram represents the modelling method used by NFM. For each variable, quantitative points are joined in a qualitative manner using fuzzy-logic. The result is seen as a curve. When summated in series and interpreted, an output is produced, in this case; time to tumour progression (months). See text for translation of case 1.

[6] M.K.B. Parmar, L.S. Freedman, T.B. Hargreave and D.A. Tolley, Prognostic factors for recurrence and follow up policies in the treatment of superficial bladder cancer: interim report from the British Medical Research Council subgroup on superficial bladder cancer (Urological cancer working party). *Journal of Urology,* 142:284-288, 1989.

[7] H.B. Burke, P.H. Goodman, D.B. Rosen, et al., Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer,* 79:857-862, 1997.

[8] S.S. Cross, R.F. Harrison, R.L. Kennedy, Introduction to neural networks. *Lancet*, 346:1075-1079, 1995.

[9] J.W.F. Catto, A. Azzouzi, I. Rehman, et al. Promoter hypermethylation is associated with tumor location, stage and subsequent progression in transitional cell carcinoma. *Journal of Oncology*, 2005 (in press).

[10] W.G. Baxt, Application of artificial neural networks to clinical medicine. *Lancet,* 346:1135-1138, 1995.

[11] K.N. Qureshi, R.N.G. Naguib, F.C. Hamdy, D.E. Neal and J.K. Mellon, Neural network analysis of clinicopathological and molecular markers in bladder cancer. *Journal of Urology,* 163:630-633, 2000.

[12] C.M. Bishop. Neural networks for pattern recognition. Oxford: Clarenden Press, 1995.

[13] M. Chen and D.A. Linkens, A systematic neuro-fuzzy modelling framework with application to material property prediction. IEEE Trans SMC. Part B: Cybernetics. 31(5):781-790, 2001.

[14] D.J. Sargent, Comparison of artificial neural networks with other statistical approaches. *Cancer*, 91:1636-1642, 2001.

[15] C. Porter, C. O'Donnel, D.E. Crawford, et al., Artificial neural network model to predict biochemical failure after radical prostatectomy. *Molecular Urology,* 5(4):159-162, 2001.

[16] S.R. Potter, C. Miller, L.A. Mangold, et al., Genetically engineered neural networks for predicting prostate cancer progression after radical prostatectomy. *Urology*, 54:791-795, 1999.

[17] A. Tewari, M. Issa, R. El-Galley, et al. Genetic adaptive neural network to predict biochemical failure after radical prostatectomy: a multi institutional study. *Molecular Urology*, 5(4):163-169, 2001.

[18] A.M. Ziada, T.C. Lisle, P.B. Snow, R.F. Levine, C. Miller and D.E. Crawford, Impact of different variables on the outcome of patients with clinically confined prostate cancer. *Cancer,* 91:1653-1660, 2001.

[19] H. Seker, M.O. Odetayo, D. Petrovic, et al., Assessment of nodal involvement and survival analysis in breast cancer patients using image cytometric data: statistical, neural network and fuzzy approaches. *Anticancer Research,* 22(1A):433-438, 2002.