PRELIMINARY ARTIFICIAL NEURAL NETWORK ANALYSIS OF SELDI MASS SPECTROMETRY DATA FOR THE CLASSIFICATION OF MELANOMA TISSUE.

LEE J LANCASHIRE[1]
SHAHID MIAN[1]
ROBERT C REES[1]
GRAHAM R BALL[1]*

[1]*The Nottingham Trent University, School of Science, Clifton Campus, Clifton Lane, Nottingham, NG11 8NS, UK.*
Correspondence author: graham.balls@ntu.ac.uk

Abstract: Over recent years studies have shown an increasing application of bioinformatics tools, and in particular, artificial intelligence techniques such as artificial neural networks (ANNs) for biological problems. Proteomic techniques such as SELDI-MS (Surface Enhanced Laser Desorption/Ionization Mass Spectrometry) may be used to distinguish patterns derived from diseased tissue used to identify biomarkers representative of a certain pathological state. This paper describes research building upon studies by Ball *et al.* (2002) in identifying potential biomarkers using ANNs for the analysis of mass spectrometry data from melanoma tissue. This approach utilised ANNs to model for 72 melanoma tissue samples (36 stage 1 and 36 stage 4) to identify ions as potential biomarkers and any possible interactions between these. Preliminary results have shown that approximately 20,000 inputs can be screened to just 20 molecular ions which are capable of accurately predicting tumour grade. Using the additive approach described by Ball *et al.* (2002) individual molecular ions were used to predict tumour grade and model performance evaluated using a receiver operating characteristic (ROC) curve. The accuracy of the model with 1 ion was 58.3% with a sensitivity of 63.9% and specificity of 52.8%. With a 2 ion model, ANN performance increased to an accuracy of 70.8%, sensitivity of 66.7% and specificity of 75%. With a 3 and 4 ion model, the accuracy increased yet further to 77.8 and 83.3%, sensitivity to 72.2 and 77.8% and specificity of 83.3 and 88.9% respectively. Preliminary findings indicate the ANN approaches adopted allow optimisation and determination of the minimum number of ions (derived from SELDI-MS data) which can successfully predict tumour grade. This work continues so that these key ions may be determined in order to identify if they have an important role in tumour progression from low to high grade.

keywords: Artificial neural networks, methodologies, models and algorithms.

1. INTRODUCTION

Melanoma is a form of skin cancer which is difficult to treat if not detected early. If however, it is detected in its early stages, survival rate is promising. Therefore, new technologies need to be developed which either (i) detect the disease at an early stage, so that it can be treated before it progresses or (ii) identify biomarkers representative of a given pathological state, which may in turn be used in the development of novel treatments.

Mass spectrometry is an important tool which is used in linking proteins to their genes [Yates, 1998]. SELDI-MS is capable of rapidly analysing samples containing vast amounts of proteins with excellent reproducibility. It can be used in generating patterns that these masses of proteins produce, and therefore is useful in showing the differences between these patterns when the proteins are being expressed in different tissues, such as differences between tissues during various stages of disease. Techniques such as this may be used to search for biomarkers associated with tumour progression and/or used in the early detection of the disease. Due to the vast amount of data generated by SELDI-MS, the development of robust computer algorithms is an absolute necessity [Ball *et al.,* 2002]. For this reason, this study involved using artificial neural networks to analyse SELDI-MS data from melanoma tissue.

ANNs are presently being utilised more than any other learning tool in biotechnology, in the modelling of complex data. This is particularly true in the field of cancer [Almeida, 2002]. Examples of their uses have been shown in prostatic cancer [Porter *et al.,* 2002; Batuello *et al.,* 2001], cervical cancer [Mango, 1998], lung cancer [Zhou *et al.,* 2002], ovarian cancer [Petricoin *et al.,* 2002] and breast cancer [Abbass, 2002; Simpson *et al.,* 1995], where ANNs have been shown to perform significantly better than physicians in the diagnosis of malignant and benign calcifications on mammograms [Markopoulos *et al.,* 2001]. ANNs have also been used in numerous other fields such

Proceedings 17th European Simulation Multiconference
(c) SCS Europe BVBA, 2003 ISBN 3-936150-25-7

as the prediction of rehospitalization in patients suffering from strokes [Ottenbacher *et al,* 2001], determining progression of glaucoma [Lin *et al.,* 2002], classification of bacterial growth [Hajmeer and Basheer*,* 2003], identifying factors which modify the responses of plant species to ozone [Balls *et al.,* 1996] and detecting the presence of fish species in rivers [Mastrorollo *et al.,* 1997].

In this study, a multi-layer perceptron ANN with a back propagation algorithm was used to model for 72 melanoma serum samples, 36 of which were low grade (stage 1) and the remaining 36 were high grade (stage 4). The purpose of the study was to identify any ions which were important in the correct classification of tumour grade and therefore may serve as potential biomarkers representative of a specific disease state. Techniques involved using relative importance values based on the weights of trained models to rank the importance of an inputs influence on the system [Balls *et al.,* 1996] and then removing inputs of low and no importance. This results in a more generalised model being developed, which enables the additive approach described by Ball *et al.* (2002) to be used in order to identify these important ions.

2. METHODS

2.1 SELDI-MS

Tissue preparation and SELDI mass spectrometry was carried out as described previously by Ball *et al.* (2002). In summary, two sequential 10-15 µm frozen tissue sections were cut for each tumour. The first section was stained with haematoxylin and eosin in order to determine tumour grade, purity and viability. The tissue was then placed directly onto 30 µl homogenising buffer (9.5 M urea, 3% CHAPS, 1 % DTT) for 15 min at room temperature with agitation to facilitate cell lysis and protein extraction. Homogenates were frozen at a temperature of -80 °C prior to SELDI analysis. Protein 'chips' were loaded with 2 µl of 50 % acetonitrile and 5 µl of cellular homogenate and exposed to the chip surface for 10 min at room temperature in a humid environment. Homogenates were then removed and the chip surface washed three times using 10 µl of water. The surface was then dried. Chip analysis was conducted at maximum laser intensity and 'phenomic fingerprints' derived from each tumour sample (Ball *et al.,* 2002).

2.2 Optimisation Of ANN Architecture

The study used a multi-layer perceptron ANN with a back propagation algorithm and a sigmoidal transfer function [Rumelhart and McClelland*,* 1986]. The particular architecture used contained three layers, with the hidden layer containing 2 hidden nodes. Determining the number of hidden nodes is essentially a trial and error procedure and 2 were found to give the best performance for this particular data (results not shown this paper). Architectures with learning rate and momentum values between 0.1 and 0.9 were trained in order to deduce the ANN model which performed best for this data. Using the mean squared error (MSE) value as a means of measuring prediction accuracy, it was found that a learning rate of 0.1 with a momentum value of 0.5 produced the lowest value. Training was conducted upon 60% of samples, with 20% being used for test sets, and the remaining 20% used for production (validation) sets until the model reached convergence. During training, the ANN model is optimised against the test set, and then validated against the production (validation) set. Convergence was determined by a failure of the model to improve the minimum MSE on the test data for 20,000 training events.

2.3. Determination Of Important Molecular Ions

The pattern recognition process involves several distinct phases which are (i) data representation, involving the initial determination of whether the tumour grade of the tissue was stage 1 or stage 4, (ii) feature extraction, where the analysis of the weights occurs, (iii) classification, where the ANN model assigns the data into either low or high grade classes and (iv) validation, where the ANN model is tested against unseen global data.

To determine which ions had the most influence on the system in correctly predicting tumour grade, the data was first screened for noise removal. To achieve this, data within the mass range of 2-5 KDa was trained over 50 random training/test/production subsets (so that a good level of confidence could be gained) and relative importance values for each individual ion was recorded in order to rank these ions according to their influence upon predicting tumour grade. These relative importance values are calculated from the analysis of weights of the trained network, values are calculated by taking the sum of the absolute weight values leading from each input to the output.

The data was then "shifted" up 500 Da so that the input data now ranged from 2.5-5.5 KDa and then trained as above. This process was repeated and the inputs were shifted over the whole data range (up to 30 KDa) which provided a proteomic profile showing relative importance of ions over the whole data range from 2-30 KDa taking account of potential interactions between ions.

From this relative importance analysis, ions with the greatest importance were selected from the data in order to reduce the number of inputs in the model. This was achieved by selecting the top 1,000 ions with the greatest relative importance values and repeating the training process as described previously. Relative importance analysis was again used to determine the top 500 ions from these 1,000. This was again repeated to deduce the top 300, 200, 100, 50, 30 then finally the top 20 ions from the initial data set of approximately 20,000 in terms of relative importance.

The next stage involved identifying the minimum number of ions from these 20 which were capable of correctly predicting tumour grade. This was achieved using an additive approach which involves training a number of different models. Using these 20 ions, each ion was used as a single input in predicting tumour grade, and for each model, 100 random training/test/production subsets were used (a process termed bootstrapping), in order to provide a measure of confidence in the predictions made. The MSE was calculated, and the ion model with the lowest value was selected for further training. All of the remaining ions were then added sequentially to this first input creating 19 two-ion models and these were trained as before with 100 random training/test/production subsets. The model with the best performance was selected to produce a three-ion model, and then the process was repeated and the ion with the best performance was again selected to produce a four-ion model.

3. RESULTS

The data obtained from SELDI-MS were analysed for relative importance values to create a relative importance profile for all data points with a m/z value of between 2 and 30,000 Daltons. Figure 1a-c shows the mean relative importance values from the 50 sub-models which were applied to each individual model.

The next stage involved selecting the ions which were most important in predicting tumour grade in order to optimise the model. This was achieved by ranking the ions in descending order of importance and selecting the top 1,000 ions. The training procedure was repeated and the top 500 ions were selected. This was repeated again so that the top 300 ions were selected and so on until a model containing the top 20 ions of importance was found. The purpose of this was to reduce the number of ions from an initial value of approximately 20,000 molecules to just the 20 that could predict tumour grade most accurately.

In order to identify the minimum number of ions which were able to accurately distinguish between low and high grade tumours, an additive approach was used (as described in the previous section). This involved creating several models and assessing their performance with respect to the MSE value generated. Preliminary results have shown that the lowest MSE value obtained from the one-ion model was from an ion with a molecular mass of 7247 achieving a MSE of 0.235. Using a two-ion model, the error decreased to 0.205 using ions 7247 and 27867. With a three-ion model, containing the ions from the two-ion model and ion 4562, the error decreased further to 0.188. Finally, with a four ion model containing the addition of ion

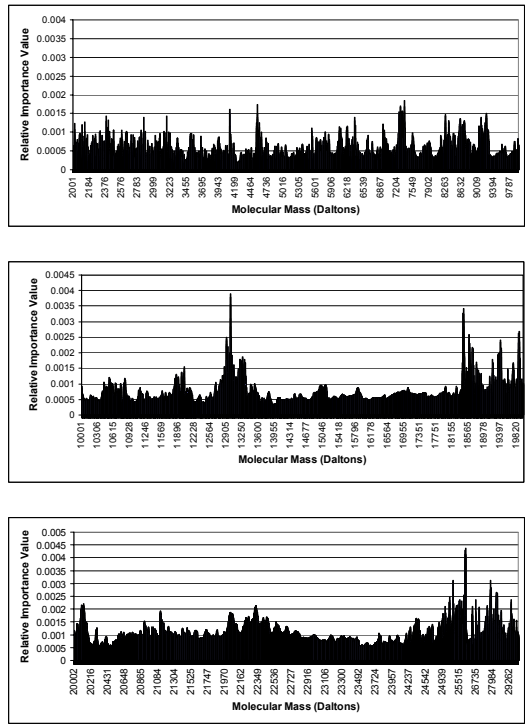

Figure 1a-c. Relative importance values for ion masses ranging from (a) 2,000-9,999 Da. (b) 10,000-19,999 Da and (c) 20,000-29,994 Da. These values illustrate a value obtained from 50 sub-models of each individual model in which different random weightings were applied to each model.

28470, the MSE value decreased further still to 0.16. Model performance from these preliminary results was then assessed using a Receiver Operating Characteristic (ROC) curve. A ROC curve determines the number of true positives (or correctly defined stage 4 melanoma tissue), true negatives (correctly defined stage 1 tissue), false positives (incorrectly defined stage 4 tissue) and false negatives (incorrectly defined stage 1 tissue). It achieves this by plotting the true positive rate against the false positive rate at different possible cutpoints (in this case, prediction errors). The curves for the one to four-ion models can be seen in Figure 2. The ROC curves for all models were compared and the results are presented in Table 1.

It is clear from Table 1 that a ROC curve provides information about several different variables. Briefly, accuracy is the overall ability of the model

Ions in model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	AUC
1 ion	58.3	63.9	52.8	57.5	59.4	0.574
2 ion	70.8	66.7	75	72.7	69.2	0.748
3 ion	77.8	72.2	83.3	81.3	75	0.809
4 ion	83.3	77.8	88.9	87.5	80	0.854

Table 1. Comparison of performances for each model used

to correctly assign the tissue samples. The sensitivity is the percentage of the stage 4 tissues correctly classified whilst the specificity is the percentage of stage 1 tissues correctly classified. The positive predictive value shows the percentage of the true positives distinguished from the false positives and the negative predictive value is the percentage of true negatives from false negatives.

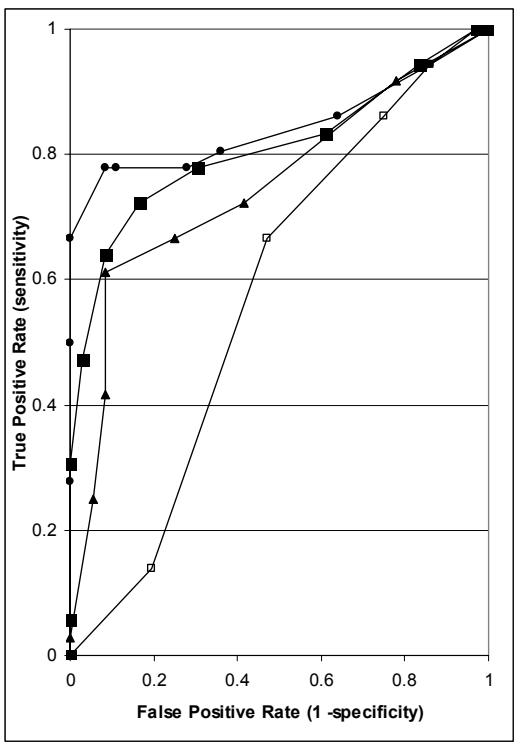

Figure 2. Diagram of ROC curves for all models. (□) shows ROC curve for one-ion model. (▲) shows ROC curve for two-ion model. (■) shows ROC curve for three-ion model. (•) shows ROC curve for four-ion model.

The area under the curve, or AUC measures discrimination, that is, the ability of the model to correctly classify those with stage 4 and stage 1 disease. A perfect ROC curve (and therefore a perfect test) has an AUC (area under the curve) value of 1, so the closer the curve follows the left hand border and then the top border of the ROC space, the more accurate the test. From the results it is clear that with increasing ions the accuracy of the model also increases. The one-ion model correctly classified 42 out of 72 tissue samples (58.3 %), with a two-ion model 51 out of 72 (70.8 %) samples were correct, using three ions, 56 of the 72 samples were assigned correctly (77.8%) and with the four-ion model, accuracy rose to 83.3 %, with 60 out of 72 samples being predicted correctly. The sensitivity and specificity of the models also showed a similar increase in performance as the accuracy. The sensitivity rose from 63.9% with the one-ion model, to 66.7 % with two-ions, a further increase was evident with the three-ion model to 72.2 % and with four-ions this rose again to 83.3 %. Meanwhile, specificity with one-ion was 52.8 % and increased by over 20 % when a second ion was added to the model (75 %). The three-ion model resulted in a specificity of 83.3 % which improved to 88.9 % for the four-ion model. The positive and negative predictive values also showed similar trends, rising from 57.5 % and 59.4 % with the one-ion model, to 87.5 % and 80 % with the four-ion model respectively. When assessing the AUC values, the one-ion model produced an AUC of 0.574 signifying a poor test. The two-ion model had an AUC of 0.748 which represents a fair test. The three and four-ion models had AUC values of 0.809 and 0.845 respectively illustrating good tests.

4. DISCUSSION

Results shown are those of preliminary work being carried out with the aim of identifying biomarkers capable of accurately predicting tumour grade from SELDI-MS data and therefore may be important in either developing novel therapies or in early detection of disease. The first stage of this study was to identify the top 20 ions (out of an initial 20,000) which were capable of predicting tumour grade. The next stage involved using an additive approach in order to determine and identify the minimum number of key ions that were capable of predicting tumour grade and thus may be important in tumour progression from low to high grade. ROC curves were generated for each model and these showed the increase in performance over the four models. The ion with the lowest error was identified as ion 7247 which predicted tumour grade with an accuracy of 58.3 %. The two-ion model contained ions 7247 and 27867 which predicted correctly 70.8 % of the time. The three-ion model consisted of ions 7247, 27867 and 4562 and classified the tissues correctly to a value of 77.8 %. Finally, the four-ion model contained ions 7247, 27867, 4562 and 28470 and performed with an accuracy of 83.3 %.

5. CONCLUSION

In conclusion, these preliminary findings show that when combining two powerful tools in SELDI-MS and ANNs, we can optimise the models so that the inputs with little or no influence upon the system may be removed in order to find the essential ions involved in the prediction of tumour grade. Although the method may be limited because the data set was relatively small due to the difficulties in obtaining tissue samples (these data sets will be expanded in future studies), the high classification accuracy of the models on truly unseen data shows that the models have generalised well enough to overcome this. Further research is ongoing and work continues on developing the models in order to conclude which, and how many ions the optimal model for the classification of tumour types from tissue contains. Once this is achieved, the next phase will involve methods for the analysis of interactions between ions within the system. Once these essential ions are identified they may be sequenced to determine their corresponding molecule/protein, which is essential in order to establish diagnostic markers.

6. REFERENCES

Abbass, H.A. 2002. "An evolutionary neural networks approach for breast cancer diagnosis". In *Artificial Intelligence in Medicine.,* 25. Pp265-281.

Almeida J.S. 2002. "Predictive non-linear modelling of complex data by artificial neural networks". In *Current Opinions in Biotechnology.* 13. Pp72-76.

Ball G. Mian S. Holding F. Allibone R.O. Lowe J. Ali S. Li G. McCardle S. Ellis I.O. Creaser C. and Rees R.C. 2002. "An integrated approach utilizing artificial neural networks and SELDI mass spectrometry for the classification of human tumours and rapid identification of potential biomarkers". In *Bioinformatics.* 18 (3). Pp395-404.

Balls G.R. Palmer-Brown D. and Sanders G.E. 1996. "Towards unravelling the complex interactions between microclimate ozone dose and ozone injury in clover". In *Water, Air and Soil Pollution.* 85. Pp1467-1472.

Batuello J, Gamito E, Crawford D. Han M. Partin A.W. McLeod D.G. and O'Donnell C. 2001. "Artificial neural network model for the assessment of lymph node spread in patients with clinically localized prostate cancer". *Urology.* 57 (3). 481-85.

Hajmeer M. and Basheer I. 2003. "Comparison of logistic regression and neural network-based classifiers for bacterial growth". In *Food Microbiology.* 20. Pp45-55.

Lin A. Hoffman D. Gaasterland D.E. and Caprioli J. 2003. "Neural networks to identify Glaucomatous visual field progression". *American Journal Ophthalmology.* 135 (1). Pp49-54.

Mango L.J. 1998. "Neural network-assisted cervical cancer screening". In *Journal of clinical ligand assay.* 21 (2). Pp203-207.

Markopoulos C. Kouskos E. Koufopoulos K. Kyriajou V. and Gogas J. 2001. "Use of artificial neural networks (computer analysis) in the diagnosis of microcalcifications on mammography". In *European Journal of Radiology.* 39, Pp60-65.

Mastrorillo S. Lek S. Dauba F. and Belaud A. 1997. "The use of artificial neural networks to predict the presence of small-bodied fish in a river". In *Freshwater Biology.* 38. Pp237-246.

Ottenbacher K.J. Smith P.M. Illig S.B. Linn R.T. Fielder R.C. and Granger C.V. 2001. "Comparison of logistic regression and neural networks to predict rehospitalization in patients with stroke". In *Journal of Clinical Epidemiology* 54. Pp1159-1165.

Petricoin E.F. Ardekani A.M. Hitt B.A. Levine P.J. Fusaro V.A. Steinberg S.M. Mills G.B. Simone C. Fishman D.A. Kohn E.C. and Liotta L.A. 2002. "Use of proteomic patterns in serum to identify ovarian cancer". In *The Lancet.* 359. Pp.572-577.

Porter C.R. O'Donnell C. Crawford D.E. Gamito E.J. Sentizimary B. De Rosalia A. and Tewari A. 2002. "Predicting the outcome of prostate biopsy in a racially diverse population: A prospective study". In *Urology.* 60 (5). Pp831-835.

Rumelhart D.E. and McClelland J.L. 1986. "Parallel distribution processing: Explorations in the microstructure of cognition". Volume 1:Foundations. MIT Press. Cambridge, MA, USA.

Yates J.R. 1998. "Mass spectrometry and the age of the proteome". In *Journal of Mass Spectrometry.* 33. Pp1-19.

Simpson H.W. McArdle C. Pauson A.W. Hume P. Turkes A. and Griffiths K. 1995. "A non-invasive test for the pre-cancerous breast". In *European Journal of Cancer.* 31A (11). Pp1768-1772.

Zhou Z. Jiang Y. Yang Y. and Chen S. 2002. "Lung cancer cell identification based on artificial neural network ensembles". In *Artificial Intelligence in Medicine,24.*Pp25-31.