Intelligent Systems for Cancer Prediction and Diagnosis

Tawfeeq Al-Kanhal[1], Yazan Otoum[2]
[1]College of Engineering, Majma'ah University
King Fahed Road, Majma'ah City, KSA
t.alkanhal@mu.edu.sa
[2]College of Computer and Information Sciences, Majma'ah University
King Fahed Road, Majma'ah City, KSA
y.otoum@mu.edu.sa

Extended Abstract

Our knowledge of genomic medicine has increased at an astonishing pace in recent years. As a result of advances in genomics, proteomics and molecular pathology many cancer biomarkers with potential clinical value have been identified. Besides, recent refinements of experimental technique demonstrate that array-based assays, generating huge datasets and finding numerous potentially useful molecules, are becoming widespread. In the present work, artificial neural network (ANN) system is used to simulate a huge database of genes to predict and diagnose the cancer disease. Experimental data collected were tested with artificial neural network technique.

The use of molecular biomarkers for cancer staging and personalization of therapy could significantly improve patient care. However, the ideal biomarker should be sensitive, specific, cost-effective, fast and robust against inter-operational and inter-institutional variability. It must also demonstrate clinical value beyond those of the other types of information that are already available at the time of diagnosis. Consequently, very few biomarkers have been incorporated into routine clinical practice. At a systems level, most of the analysis and modelling of genetic experimental data has utilised linear technologies. However, at all levels of physiology, non-linear events are paramount. Recent advances in intelligent systems technology are enabling non-linear phenomena to be captured quite frequently.

In this context, the major aim of this research is to use intelligent systems techniques to identify a group of biomarkers that not only optimise the diagnosis and behaviour prediction of bladder cancer, but also improve patients care.

At a more specific level, it is recognized that numerous molecular events in cancer occur at the epigenetic rather than genetic level. In normal cells, the cytosine-guanine dinucleotides (CpG islands) found in genomic promoter regions are unmethylated for the majority of genes and this allows gene expression. In cancer, loss of control of DNA methylation results in regional CpG hypermethylation and associated tumour suppressor gene silencing. As high levels of aberrant hypermethylation occur early in the natural history of cancer and are most common in aggressive tumours, this molecular event appears as an excellent clinical biomarker for disease progression [1- 10]. For example, changes in DNA hypermethylation can be detected in the normal bladder urothelium of patients with invasive Urothelial Cell Carcinoma (UCC) and are found to predict disease biology [11].

Whilst cells with DNA methylation can be detected in the urine and used to diagnose UCC, many epigenetic events are not cancer specific and are found within normal aging cells [12]. Thus, the current challenge facing epigenetic researchers is the identification of cancer-specific genes that will allow the application of this knowledge to clinical practice. One cancer-specific methylated gene has been found in prostate cancer (GSTP1). When compared, GSTP1 methylation was superior to expert histopathological examination at recognising false negative prostate biopsies for cancer detection [13].

In order to identify methylated genes that are either cancer or behaviour specific, microarray technology is the best tool. These arrays simultaneously study 12,000 CpG islands spread over the entire genome and identify those islands differentially methylated between comparative tissues (e.g. cancer and normal DNA). Whilst the cost of these experiments is high and so the sample sizes are modest (in comparison with studies using less advanced technologies), they generate huge datasets. For example, studying 12,000 CpG islands in 100 samples and analysing the results against 5 different clinical parameters results in 6 million experiments/comparisons if each island is analysed in isolation. Furthermore,

combinational island analysis is probably more accurate in terms of biology, but the datasets can be prohibitive whilst using current statistical methods.

This research work is based on studying epigenetic data for the diagnosis and progression of UCC. The plan is to rank, using a systems-based approach, 12,000 CpG islands/genes according to their association with various clinico-pathological criteria, including tumour behaviour.

References

[1] M. F. Abbod, N. Mort and D. A. Linkens, "Fuzzy Methods for the Modelling and Automatic Control of Material Finishing," *COMEDEM'95*, Queens University, Canada, pp 231-236, 1995.

[2] M. F. Abbod, J. W. F Catto, M. Chen, D. A. Linkens, and F. C. Hamdy, "Artificial Intelligence for the Prediction of Bladder Cancer," *Journal of Biomedical Engineering, Applications, Basis, and Communication*, vol. 16, no. 2, pp. 7-16, 2004.

[3] M. F. Abbod, C. M. Sellars, D. A. Linkens and Q. Zhu, "Validation and Generalisation of Hybrid Models for Flow Stress and Recrystallisation Behaviour of Aluminium Alloys," *Materials Science and Engineering A*, vol. 395, no. 1-2, pp 35-46, 2005.

[4] M. Mahfouf, J. Tenner, D. A. Linkens and M. F. Abbod, "Optimal Design of Alloy Steels Using Evolutionary Computing," *4th Int Conf on Knowledge-Based Intelligent Engineering Systems & Allied Technologies*, Brighton, University of Brighton, 2000.

[5] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," *Proc. IEEE int'l conf. on neural networks*, vol. IV, pp. 1942-1948, IEEE service centre, Piscataway, NJ, 1995.

[6] M. Jamei, M. Mahfouf, D. A. Linkens, "Optimal steel design using multi-objective optimisation algorithm," *International Conference on Metal Fabrication and Welding Technology (METFAB – 2003) and European Symposium on Manufacturing and Modelling of Fabricated Structural Components (MMFSC)*, Nottingham Royal Moat House, Nottingham, 2003.

[7] J. W. F. Catto, D. A. Linkens, M. F. Abbod, M. Chen, J. L. Burton, K. Feeley, M. Meuth and C. F. Hamdy, "Artificial Intelligence in Predicting Bladder Cancer Outcome: A Comparison of Neuro-Fuzzy Modelling and Artificial Neural Networks," *Clinical Cancer Research*, vol. 9, pp. 4172-4177, 2003.

[8] J. W. F. Catto, M. F. Abbod, D. A. Linkens and C. F. Hamdy, "Neuro-fuzzy modelling: An Accurate and Interpretable Method for Predicting Bladder Cancer Progression," *The Journal of Urology*, vol. 175, pp. 474-479, 2006.

[9] J. W. F. Catto, A. R. Azzouzi, I. Rehman, K. M. Feeley, S. S. Cross, N. Amira, G. Fromont, M. Sibony, O. Cussenot, M. Meuth and F. C. Hamdy, "Promoter hypermethylation is associated with tumor location, stage and subsequent progression in transrtional cell carcinoma," *J Clin Oncol*, vol. 23, pp. 2903-2910, 2005.

[10] M. V. Brock, M. Gou, Y. Akiyama, A. Muller, T. T. Wu, E. Montgomery, M. Deasel, P. Germonpre, L. Rubinson, R. F. Heitmiller, S. C. Yang, A. A. Forastiere, S. B. Baylin and J G. Herman, "Prognostic importance of promoter hypermethylation of multiple genes in esophageal adenocarcinoma," *Clin Cancer Res*, vol. 9, no. 8, 2003, pp. 2912-2919.

[11] D. Dhawan, F. C. Hamdy, J. Patterson, I. Rehman, K. Feeley, M. Meuth and J. W. F. Catto, "Promoter methylation in pre-malignant and malignant urothelial carcinoma: a feature of the invasive tumor phenotype," *J Pathol*, vol. 209, no. 3, pp. 336-343, 2006.

[12] D. R. Yates, I. Rehman, M. Meuth, S. S. Cross, F. C. Hamdy and J. W. F. Catto, "Methylational urinalysis: a prospective study of bladder cancer patients and age stratified benign controls," *Oncogene*, vol. 15, pp. 1984-1988, 2006.

[13] S. V. Harden, Z. Guo, J. I. Epstein, D. Sidransky, "Quantitative GSTP1 methylation clearly distinguishes benign prostatic tissue and limited prostate adenocarcinoma," *J Urol*, vol 169, no 3, pp 1138-1142, 2003.