Introduction and investigation of the Top scoring pair method by detecting the efficiency of drug in breast cancer

Saeed Rasekhi Dehkordi, Ebrahim Hajizadeh, Ali Mohammadian
Department of biostatistics, Tarbiat Modares University, saeid.rasekhi@modares.ac.ir, Iran
Department of biostatistics, Tarbiat Modares University, hajizadeh@modares.ac.ir, Iran
Department of biotechnology. Tarbiat Modares University, a.mohammadian@modares.ac.ir, Iran

Abstract: Gene-based cancer classification comprising multiple discriminatory molecular markers is an emerging trend. In this paper, we introduce a new approach for classifying binary disease states, referred to as the top-scoring pair(s) (TSP)[1]. This method focuses on detecting small pairs of genes in which the comparison of their expression values leads to class discrimination. Accurate statistical inference from such data is difficult due to small number of samples, typically tens, relative to the large number of genes, typically thousands in microarray experiments. Moreover, conventional methods from machine learning lead to decisions which are usually very difficult to interpret in simple or biologically meaningful terms. In contrast, the TSP classifier provides decision rules which: i) involve few genes and only relative expression values; ii) are both accurate and transparent; and iii)[2] provide specific hypotheses for follow-up studies. We investigated this method by one of the most challenging issues in medical science which is drug efficiency in cancer cases. One of the important factors in the improvement of a patient, in particular cancer disease, is how effective the drug is. In this study the aim was finding effective biomarkers in breast cancer.[3] The database include 508 breast cancer patients in which 169 patients was predicted to be sensitive to drug and in 339 cases it was insensitive. The accuracy of this method is %70. Furthermore, to better perception sensitivity and specificity was calculated, which are 0.74 and 0.71 respectively. **Keywords:** microarray data; class prediction; gene comparisons

References

[1] Geman D, d'Avignon C, Naiman DQ, Winslow RL. “Classifying gene expression profiles from pairwise mrna comparisons” Statistical Applications in Genetics and Molecular Biology. 3(2004)

[2] Sitan Yang, Daniel Q. Naiman, “Multiclass cancer classification based on gene expression comparison” Statistical Applications in Genetics and Molecular Biology. 3(2014)

[3]Matthew D. Burstein, Anna Tsimelzon, Graham M. Poage, Kyle R. Covington, Alejandro Contreras, Suzanne A.W. Fuqua, Michelle I Savage, C. Kent Osborne, Susan G. Hilsenbeck, Jenny C. Chang, Gordon B. Mills, Ching C. Lau, and Powel H. Brown.” Comprehensive Genomic Analysis Identifies Novel Subtypes and Targets of Triple-negative Breast Cancer” Clin Cancer Res. 4(2016)