

---

## Editorial

---

### Sun Kim

School of Informatics and Computing,  
Indiana University,  
Bloomington, Indiana, USA  
E-mail: sunkim2@indiana.edu

**Biographical notes:** Sun Kim is an Associate Professor and Director of Center for Bioinformatics Research in the School of Informatics and Computing at Indiana University, Bloomington, IN USA. He received BS and MS and PhD in Computer Science from Seoul National University, Korea Advanced Institute of Science and Technology (KAIST) and the University of Iowa, respectively. His research interests are development and applications of machine learning algorithms to the analysis of microbial genome and cancer epigenomics data and also on the bioinformatics systems architecture in the cloud computing environments.

---

### 1 Introduction

This is a second issue of papers that were selected and extended among top papers from 233 submissions at the third IEEE International Conference on Bioinformatic and Biomedicine in 2009. The rapid accumulation of biological data requires more sophisticated, efficient computational methods at almost all stages of research in biology and medical sciences from handling and analysis of data to mining novel insights on new knowledge in biology. This special edition of the *International Journal of Data Mining and Bioinformatics* contains another set of seven papers that report how bioinformatics techniques can be used for various aspects of biological research.

The first paper by Pyon, Y.S., Li, X. and Li, J. ‘Cancer progression analysis based on ordinal relationship of cancer stages and co-expression network modularity’, proposed a multicategory logit model to identify genes that show significant correlations across multiple cancer stages. By using a prostate cancer progression data, the authors predicted a set of known prostate-cancer-related genes that included genes that were confirmed only very recently.

The second paper by Nelson, S.J., Zeng, J. and Kilbourne, J. ‘Building a standards-based and collaborative e-prescribing tool: MyRxPad’, reported MyRxPad ([rpx.nlm.nih.gov](http://rpx.nlm.nih.gov)), a prototype application that allows interaction between practitioners and patients using e-prescribing. Patients play an active role by maintaining up-to-date and accurate medication lists. Prescribers make well-informed and safe prescribing decisions based on personal medication records contributed by patients. The paper reported experience in applying RxNorm in an e-prescribing setting.

The third paper by Rojas, D., Rueda, L., Ngom, A., Hurrutia, H. and Crcamo, G. ‘Image segmentation of biofilm structures using optimal multi-level thresholding’, proposed a combination of techniques for segmentation of biofilm images through an

optimal multi-level thresholding algorithm and a set of clustering validity indices. Since manual appreciation of biofilm structures in digital images varies significantly, the segmentation method will be useful. Experiments showed that the quality of segmentation predicted by the method is comparable to that by an expert in terms of Rand Index.

The fourth paper by Nasser, S., Ranade, A.R., Sridhar, S., Haney, L., Korn, R.L., Gotway, M.B., Weiss, G.J. and Kim, S. ‘Biomarkers associated with metastasis of lung cancer to brain predict patient survival’, reported an approach that predicted biomarkers by combining validated miRNA expression values and imaging features. The methods used in the approach included miRNA expression profiling, evaluation of normalisation techniques and combination of miRNA with imaging features FDG-PET/CT and CT Scan. The predicted biomarkers were validated on an independent data set.

The fifth paper by Lv, D. and Goodwine, B. ‘Pancreas modelling by a deterministic optimisation method’, presented a mechanistic pancreas model of insulin dynamics using experimental physiological data. Intravenous Glucose Tolerance Test (IVGTT) data from the literature was used for the study. The model parameters were predicted by using a deterministic optimisation method called DIRECT (DIviding RECTangles).

The sixth paper by O’Connor, M.J., Hauser, A.E., Haberman, A.M. and Kleinstein, S.H. ‘Activated germinal centre B cells undergo directed migration’, used statistical approaches to show that activated B cells move in a directed manner that sharply contrasts with the behaviour of naïve B cells for the study of affinity maturation. The data used for the study was a data set collected by using two-photon microscopy that tracked individual cells moving within germinal centres in living animals.

The final paper by Alshalalfa, M., Naji, G., Qabaja, A., Alhajj, R. and Rokne, J. ‘Combining multiple perspective as intelligent agents into robust approach for biomarker detection in gene expression data’, reported a computational framework that incorporates different perspectives each realised by an agent. Each agent applies its own method to analyze the gene expression data and to come up with candidate genes as potential cancer biomarkers. Gene enrichment, protein interaction, and miRNA regulation were given weights and they were used to confirm the discoveries by the major agents. The authors conducted experiments on two data sets and obtained prediction result with a high classification rate.

Together with six papers in the previous volume, the seven papers in this volume provide scientists a glimpse on how bioinformatics techniques are used for research in biology and medical sciences and I hope the papers can boost up more extensive use of bioinformatics techniques for research in biology and medical sciences.