Chapter 2.12
Integration of Clinical and Genomic Data for Decision Support in Cancer

Yorgos Goletsis
University of Ioannina, Greece

Themis P. Exarchos
University of Ioannina, Greece

Nikolaos Giannakeas
University of Ioannina, Greece

Markos G. Tsipouras
University of Ioannina, Greece

Dimitrios I. Fotiadis
University of Ioannina, Greece, Michaelideion Cardiology Center, Greece & Biomedical Research Institute, Greece

INTRODUCTION

Computer aided medical diagnosis is one of the most important research fields in biomedical engineering. Most of the efforts made focus on diagnosis based on clinical features. The latest breakthroughs of the technology in the biomolecular sciences are a direct cause of the explosive growth of biological data available to the scientific community. New technologies allow for high volume affordable production and collection of information on biological sequences, gene expression levels and proteins structure on almost every aspect of the molecular architecture of living organisms. For this reason, bioinformatics is asked to provide tools for biological information processing, representing today's key in understanding the molecular basis of physiological and pathological genotypes. The exploitation of bioinformatics for medical diagnosis appears as an emerging field for the integration of clinical and genomic features, maximizing the information regarding the patient's health status and the quality of the computer aided diagnosis.

DOI: 10.4018/978-1-60960-561-2.ch212

Cancer is one of the prominent domains where this integration is expected to bring significant achievements. As genetic features play a significant role in the metabolism and the function of the cells, the integration of genetic information (proteomics-genomics) to cancer-related decision support is now perceived by many not as a future trend but rather as a demanding need. The usual patient management in cancer treatment involves several, usually iterative, steps consisting of diagnosis, staging, treatment selection, and prognosis. As the patient is usually asked to perform new examinations, diagnosis and staging status can change over time. On the other hand, treatment selection and prognosis depend on the available findings, response to previous treatment plan and, of course, clinical guidelines. The integration of these evolving and changing data into clinical decision is a hard task which makes fully personalised treatment plan almost impossible. The use of clinical decision support systems (CDSSs) can assist in the processing of the available information and provide accurate staging, personalised treatment selection, and prognosis. The development of electronic patient records and of technologies that produce and collect biological information have led to a plethora of data characterizing a specific patient. Although this might seem beneficial, it can lead to confusion and weakness concerning the data management. The integration of the patient data (quantitative) that are hard to be processed by a human decision maker (the clinician) further imposes the use of CDSSs in personalized medical care (Louie, Mork, Martin-Sanchez, Halevy, & Tarczy-Hornoch, 2007). The future vision—but current need—will not include generic treatment plans according to some naive reasoning, but totally personalised treatment based on the clinicogenomic profile of the patient.

In this article, we address decision support for cancer by exploiting clinical data and identifying mutations on tumour suppressor genes. The goal is to perform data integration between medicine and molecular biology by developing a framework where clinical and genomic features are appropriately combined in order to handle cancer diseases. The constitution of such a decision support system is based on (a) cancer clinical data and (b) biological information that is derived from genomic sources. Through this integration, real time conclusions can be drawn for early diagnosis, staging and more effective cancer treatment.

BACKGROUND

Clinical Decision Support Systems are active knowledge systems which use two or more items of patient data to generate case-specific advice (Fotiadis, Goletsis, Likas, & Papadopoulos, 2006). CDSSs are used to enhance diagnostic efforts and include computer based programs that, based on information entered by the clinician, provide extensive differential diagnosis, staging (if possible), treatment, follow-up, and so forth. CDSSs consist of an inference engine that is used to associate the input variables with the target outcome. This inference engine can be developed based either on explicit medical knowledge, expressed in a set of rules (knowledge based systems) or on data driven techniques, such as machine learning (Mitchel, 2006) and data mining (intelligent systems) (Tan, Steinbach, & Kumar, 2005). CDSSs require the input of patient-specific clinical variables (medical data) and as a result provide patient specific recommendation.

Medical data are observations regarding a patient, including demographic details (i.e., age, sex), medical history (i.e., diabetes, obesity), laboratory examinations (e.g., creatinine, triglyceride), biomedical signals (ECG, EMG), medical images (i.e., MRI, CT), and so forth. Demographic details, medical history, and laboratory data are the most easily obtained and recorded and, therefore, most commonly included in electronic patient records. On the other hand, biomedical signals and medical images require more effort in order to be acquired