A data mining system for chronic diseases surveillance for primary care unit in Taiwan

HM Hua¹, T. Lee².
Graduate school of Gerontic Technology and Service Management, Nan Kai Institute of Technology, Caotun, Nantou, Taiwan. ¹
Doctor Lee Family Medicine Clinic, Tamshui, Taipei Hsien, Taiwan. ²

Abstract—Chronic diseases affect a large proportion among elderly and cause a major public health burden in Taiwan. Complications of hyperlipidemia, hypertension, triglyceride and diabetes highly increase the risks of medical treatment and long-term care for the elderly. In this study, we provide a data mining system for chronic diseases surveillance to assist clinics doctors in medical treatment concerns for their patients. It is not only able to reduce the cost of national health insurance but also to afford the better quality of life for elderly people.

I. INTRODUCTION

CHRONIC diseases affect a large proportion among elderly and cause a major public health burden in Taiwan [1] (Figure 1). Elderly with chronic disease is not only to increase death but also to raise health insurance cost. Complications of hyperlipidemia, hypertension, triglyceride and diabetes highly increase the risks of medical treatment and long-term care for the elderly. In this study, we provide a data mining system for chronic diseases surveillance to assist clinics doctors in medical treatment concerns for their patients. It is not only able to reduce the cost of national health insurance but also to afford the better quality of life for elderly people.

II. METHODS

Samples and Data Collection

Medical patient data in electronic form in database in this study has been obtained from a local clinic in Taipei, Taiwan from 1999 to 2007. Data also consists of all the information required to document the physician’s diagnoses and the procedures performed. Family tree and history also correspond with patients. The central role of data mining uses database and statistical methods including structured patient information with diagnosis, drugs and biochemical tests or physical examinations.

We infer the values of several medical outcomes, described by a set of variables of interest. Data collection of such variables include whether a patient has a particular disease, whether a patient has received a certain type of medication, lab recordings for blood glucose, whether a patient has specific contraindications for a class of medication (Figure 2).

Figure 1 The National Health Insurance Statistics—by diseases, 2004 data from Bureau of National Health Insurance, Taiwan

Classification of Diseases

To structured information about patient demographics, the domain knowledge of chronic diseases is classified according to the internationally accepted standards, ICD-9 (International Classification of Diseases, 9th Revision [2]. Many of the criteria used to determine if a patient is eligible for or treated according to a particular guideline, are based upon diagnostic information.

Machine Learning Approach

Many studies have shown that the clinical accuracy of ICD codes is only 60%-80% [3]. Thus, we provide a machine learning approach to data mining and knowledge discovery based on the induction of clinical classification, association and sequence rules [4].

Chronic diseases require multiple decisions concerning medications. A decision tree is regarded as a hierarchically organized set of clinical rules. Association rule is mainly for discovering chronic diseases possible synchronize. Both genetic algorithm and case base reasoning in hybrid data mining technology are used to acquire of knowledge from database [8]-[12]. These techniques also map personal, familial, and professional values to meet clinical diagnosis
requirement and long-term care decisions (Figure 3).

Our system has also built up family medical history records. It provides interdisciplinary medical and health service in prevention, emergency and chronic disease treatment. Through medical treatment analysis, clinical doctors at primary care unit are able to build up a good long term relationship with chronic diseases patients. Chronic diseases patients obtain better quality of clinical care. Clinics, doctors and patients all benefit from our chronic diseases surveillance system.

In the more distant future, drugs analysis will be added in to measure the clinical effects associated with the provision of medication therapy management services to cut down national health insurance cost.

**Auto-remind System**

Traditional medical treatment on patients relies upon manual conversion of data by medical experts. However, it is expensive, time consuming, and is only possible for a small subset of patients or at institutions with a strong research focus. Based on the data statistic and analysis, we count risk factors related with chronic diseases and patients. An auto-remind mechanism on multi-source data based on prior medical knowledge is processed with observations to assist clinics doctors in medical treatment concerns for their patients.

III. RESULTS AND DISCUSSION

Chronic diseases treatment requires long term caring. Chronic medical data is also highly complex and difficult to analyze. Lack of cost and staff, the primary care currently focuses on non-serious diseases but requiring urgent treatment. However, our chronic diseases surveillance is built on structured information in the clinical repository, typically as unstructured free text in patient history and physicals, discharge summaries, progress notes, radiology reports, etc. To treat and monitor these patients with chronic diseases complications, such as diabetic patients with hyperlipidemia or hyperlipidemia patients with triglyceride, highly improve the quality of clinical care.

Chronic disease caring relies on patients’ taking the initiative for return treatment. Our system is designed to provide the patients list according to their return treatment dates. Medical staffs take the initiative to remind patients for their return treatment. It highly improves chronic diseases treatment.

Clinical records and medical treatment analysis give a practical controlled picture to both doctors and chronic diseases patients. These analyses give great assistance to the medical staff in monitoring the patients’ symptoms.

Hypertension is an important factor among many other cardiovascular diseases. However, other risk factors such as diabetes, high cholesterol, smoking, sex, and family history etc. should also be assessed and analyzed in order to monitor the level of danger of other heart and blood related disorders. In principle, the high risk patients (i.e. the high risk groups) should have priority to be the monitored subjects.

**REFERENCES**


