

**The Dynamics of Health Care:
Tracing the Diabetic Healthcare Journey**

A dissertation submitted to The University of Manchester for the
degree of Master of Science in the Faculty of Engineering and
Physical Sciences

2012

Parivash Ashrafi

SCHOOL OF COMPUTER SCIENCE

Abstract

A large amount of healthcare data is captured in form of the Electronic Patient Records (EPR). In primary care, EPR data is captured by general Practitioners (GP). These health records contain information about the patients' history of diseases, treatments and medications. Almost all GPs in UK use "Read code" to record the patients' information in the database. This data is a valuable source of information and could be used for data management and further analysis about the diseases interactions and drugs.

In this study, we have access to the data set from primary care in Salford. The data set is in the form of Read codes and consists of 11 million records from 13500 anonymised patients. In order to visualise this large volume of data, techniques based around semantic similarity and Principal Component Analysis (PCA) have been used to take this data described as terms from ontology, and map it to a low dimensional vector space. These data then have been examined over time to follow the patients' healthcare before and after the diagnosis of diabetes.

Diabetes has substantial effects on the health care system, due to its connection with many other diseases. We have examined patient records, before and after the diagnosis with diabetes, to explore the dynamics of the diabetes healthcare journey. This is being done, by mapping the patient data into a low dimensional vector space and then applying the standard machine learning strategies to cluster and classify patients' data and finding the missed opportunities.

This study, which is performed by the collaboration with AstraZeneca Company, explores to what extent, these trajectory patterns provide useful information on patients' outcome. Looking at the trajectory before diabetes diagnosis, it is evident that there are significant missed opportunities to identify diabetic patients. In addition, the patterns after diabetes identification, gives insight about some of the co-morbidities that goes with diabetes.