Chapter 12
Mining Electronic Health Records to Guide and Support Clinical Decision Support Systems

Jitendra Jonnagaddala
University of New South Wales, Australia

Pradeep Ray
University of New South Wales, Australia

Hong-Jie Dai
National Taitung University, Taiwan

Siaw-Teng Liaw
University of New South Wales, Australia

ABSTRACT

Clinical decision support systems require well-designed electronic health record (EHR) systems and vice versa. The data stored or captured in EHRs are diverse and include demographics, billing, medications, and laboratory reports; and can be categorized as structured, semi-structured and unstructured data. Various data and text mining techniques have been used to extract these data from EHRs for use in decision support, quality improvement and research. Mining EHRs has been used to identify cohorts, correlated phenotypes in genome-wide association studies, disease correlations and risk factors, drug-drug interactions, and to improve health services. However, mining EHR data is a challenge with many issues and barriers. The aim of this chapter is to discuss how data and text mining techniques may guide and support the building of improved clinical decision support systems.

INTRODUCTION

In the last decade, a rapid adoption of electronic health record (EHR) systems has been observed in the healthcare sector in countries like Australia, Canada, UK and USA. The major purpose of EHR systems is to improve patient outcome by seamlessly exchanging patient data between healthcare providers. At the same time, EHR systems are incrementally growing in size and diversity. With the surge of new computational techniques and technologies, healthcare professionals, researchers, administrators and educators are keen to use patient data collected in EHR systems for multiple purposes. One major example is their use in clinical decision support.

DOI: 10.4018/978-1-4666-9432-3.ch012

Clinical decision support systems (CDSS) are defined as "any computer program designed to help healthcare professionals make clinical decisions" (Shortliffe, 1987) and many types of CDSS exist. The most common type includes knowledge portals and alert systems (Musen, Middleton, & Greenes, 2014). CDSS are mainly targeted towards clinicians but also target others, for example quality administrators, managers and researchers. CDSS may exist independently or work in conjunction with other hospital information systems. The latter have greater impact on patient outcomes over stand-alone CDSS (Kawamoto, Houlihan, Balas, & Lobach, 2005).

In the past, CDSS have been built on knowledge from the literature, guidelines and randomized controlled clinical trials. However, most of this knowledge applies to targeted cohorts or patients. CDSS developers and researchers can also leverage the data routinely collected on patients in EHRs. EHR data from recent patients can be exploited with data and text mining techniques to enhance the development and use of CDSS. These techniques are useful to extract information from EHRs, improve knowledge and combine the newly gained knowledge with existing information or knowledge. Despite these advantages, data and text mining techniques still present new challenges, such as the complex evolution of the amount and variety of clinical data in health information systems, poor data quality, incomplete data and mining EHR data without compromising the privacy of the patients (Botsis, Hartvigsen, Chen, & Weng, 2010).

The aim of this chapter is to provide an overview of how EHR data can be used to guide and support good clinical decision support systems, drawing on recently published literature. In the following sections, a background review on EHR and CDSS is followed by an overview of how basic data and text mining techniques are applied in a clinical domain. Cases of popular CDSS usage where EHR data are mined will also be discussed, as well as emerging issues in mining EHR data and their impact on CDSS.

BACKGROUND

In a recent perspective article published in the New England Journal of Medicine (Gandhi, Zuccotti, & Lee, 2011), the authors summarized an analysis of outpatient records for 1.7 million patients in an integrated healthcare system in the USA. The analysis highlighted that 71% of the patients who had undergone splenectomy did not have it medically coded under the 'problems list' in their EHR. As a result of this missing information, only a third of the patients whose records did not mention splenectomy were given pneumococcal vaccination. More interestingly, the authors reported that an alert-and-reminders-based clinical decision support system was already in place and was not of much assistance. In this example, text mining could be employed to medically code splenectomy under the 'problems list' in an EHR. Furthermore, data mining techniques like predictive models can be built to predict other associated problems related to splenectomy.

It is evident from the example given above that CDSS need as much of the available data as possible to guide healthcare professionals in their decision making process. The underlying problem here seems to be the lack of accessibility to necessary data required for CDSS. A recent review in 2010 reported that most reminder-based CDSS had low to medium impact on decision support and very few had high impact (Shojania et al., 2010). The reasons varied depending on the healthcare setting. However, it is important to note that the impact of CDSS is not high mainly because of clinicians overriding the decision support provided by CDSS. CDSS could give better aid by leveraging on patient data holistically just like a clinician, instead of concentrating on medications, vaccinations or diagnosis only.

Related Content