Chapter 10
Semantic Integration of Patient Data for Clinical Decision Support in Breast Cancer Care

Leonardo Lezcano
University of Alcalá, Spain

Miguel-Ángel Sicilia
University of Alcalá, Spain

Eydal Rivero
University of Alcalá, Spain

ABSTRACT
Achieving semantic interoperability between heterogeneous healthcare systems and integrating clinical guidelines in the automatic decision support of healthcare institutions are two key priorities of current medical informatics. They can lead to a significant improvement on patient safety by reducing medical risks and delays in diagnosis, facilitating continuity of care and preventing life threatening adverse events. The present chapter describes a project that addresses those two priorities in the field of Breast Cancer for which effective clinical guidelines are available, as well as the clinical data to apply them. However, the deployment of semantic interoperability techniques based on clinical terminologies such as SNOMED-CT and EHR exchange models such as openEHR and HL7 is required to meaningfully combine the available data. Then data mining techniques are capable of automatically adapting the parameters of clinical guidelines to the particular conditions of each healthcare environment.

1. INTRODUCTION
Like many other fields that heavily rely on the capabilities of information and communication technologies, healthcare and biomedical environments are rapidly increasing the demand for widely accepted agreements on data, information and knowledge exchange. Such needs for compatibility or interoperability go beyond syntactical and structural issues as semantic interoperability is also required. Semantic interoperability (Stroetmann et al.,...
Semantic Integration of Patient Data for Clinical Decision Support

2009) is essential to facilitate the computerized support for alerts, workflow management and evidence-based healthcare across heterogeneous Electronic Health Record (EHR) systems.

The model of clinical archetypes supported by the CEN/ISO EN13606 standard and the openEHR foundation (Beale, 2002) provides a mechanism to express data structures in a shared and interoperable way. It has acquired considerable acceptance in the last years by allowing the definition of clinical concepts based on a common Reference Model while low level storage implementation can keep its heterogeneity across EHR systems. However, archetype languages do not provide direct support neither for clinical rules nor mappings to formal ontologies, which are both key elements of full semantic interoperability as they allow exploiting reasoning on clinical knowledge (Lezcano et al., 2011).

It has been acknowledged that the World Wide Web demands analogous capabilities to those mentioned above, leading to the development of the Semantic Web extension. The progress made in that field, regarding reasoning and knowledge representation, is combined in this chapter with EHR models in order to enhance the archetype approach and to support features that correspond to a richer level of semantic interoperability.

The present section introduces a research carried out during the last 3 years to increase patient safety and improve outcomes in the particular context of breast cancer treatment. Most of this work is part of the HEDECAMA project, which stands for “Semantic Model and Data Mining Algorithms applied to Breast Cancer Treatment in Specialized Care Centers.” The general goal is to develop a series of complementary technology components and applications in order to obtain a clinical decisions support system (CDSS) for breast cancer treatment at care centers (i.e., hospitals, oncology centers, etc.). As a result, the end-user or breast cancer specialist receives relevant information to certain situations and clinical episodes, allowing him/her to access and process in real-time all current and historical information related to the patient’s clinical case. This would be an extremely time consuming task if accomplished manually.

The developed solution is configurable and extensible as the integrated technological elements follows commonly accepted standards. They can collect and interpret a wide range of data from the heterogeneous clinical information systems currently available in healthcare institutions, and at different stages of the clinical diagnosis and treatment of breast cancer.

This chapter is structured according to Figure 1. The workflow first includes an Integration stage, based on a Data Integration Framework that is described in Section 2 and a Terminology Server introduced in Section 3. Then the Storage and Classification stage, whose main component is presented in Section 4 interacts with the Knowledge and Machine Learning components described in sections 5 and 6. A visualization stage, the features of which are not covered in the chapter, offers the results to the end-user. Although all the stages will be explained, the current chapter is oriented to put more emphasis on the Data Integration Framework, whose details are provided in the subsections of Section 2. Finally, Section 7 provides the conclusions of the research and its impact from a clinical perspective.