Chapter XXII
Grid Technologies in Epidemiology

Ignacio Blanquer
Universidad Politécnica de Valencia, Spain

Vicente Hernández
Universidad Politécnica de Valencia, Spain

ABSTRACT

Epidemiology constitutes one relevant use case for the adoption of grids for health. It combines challenges that have been traditionally addressed by grid technologies, such as managing large amounts of distributed and heterogeneous data, large scale computing and the need for integration and collaboration tools, but introduces new challenges traditionally addressed from the e-health area. The application of grid technologies to epidemiology has been concentrated in the federation of distributed repositories of data, the evaluation of computationally intensive statistical epidemiological models and the management of authorisation mechanism in virtual organisations. However, epidemiology presents important additional constraints that are not solved and harness the take-off of grid technologies. The most important problems are on the semantic integration of data, the effective management of security and privacy, the lack of exploitation models for the use of infrastructures, the instability of Quality of Service and the seamless integration of the technology on the epidemiology environment. This chapter presents an analysis of how these issues are being considered in state-of-the-art research.

INTRODUCTION

Epidemiology is defined as “the study of factors affecting the health and illness of populations, which serves as the foundation and logic of interventions made in the interest of public health and preventive medicine” (wikipedia, 2008). The epidemiology study is the main support for public
Grid Technologies in Epidemiology

health research which involves those issues that are relevant for a population as a whole, rather than the specific treatment of individuals. Epidemiology is closely related with public health, which uses epidemiology and geographic information as main tools to describe the outbreak, evolution and eradication of diseases.

The management of individuals’ health is normally structured around the concept of Electronic Health Record (EHR) (EUROREC, 2002). Notwithstanding the commonalities between epidemiology and EHR, there are several important differences which should be considered:

• Specificity and Sensibility in epidemiology studies have different relevance in individual clinical practice and global population analysis, since one erroneous or missing register could not affect a global result, but can be critical for an individual.
• Epidemiology studies are not bounded by interactivity constraints since they are not used at the delivery of health.
• Confidentiality is considered at a different level, since data is anonymised or pseudo-anonymised, requiring different (additional in some cases) techniques.
• The level of detail of the information managed is coarser, considering that the aim is to find and study general patterns.
• Along with the clinical data, information from models, clinical trials, geographic data and social information is of high importance.

However, there are several issues that are shared with EHR:

• A large part of the data comes from clinical practice, and thus it goes through the needs and problems of data coding and integration.
• Integration of data around the identity of the patient. Although data is considered at the level of global population, information coming from individuals must be integrated in the form of complete entries.
• Need for normalisation, standard coding and standard interfaces, which are even more important in the case of epidemiology, as the data will be automatically processed.

The main needs of epidemiology are the integration and processing of clinical, social and geographical data through statistical methods and models. These issues require tackling the following technological challenges.

• Integrate the information from diverse sources of medical, social and geographical data for the construction of data warehouses, knowledge bases or similar data repositories to drive the epidemiology research.
• Use epidemiology models, data mining tools, biocomputing and other tools to extract correlations, predict effects and guide public health policies.
• Ensure a correct and legal use of the information, without compromising the privacy of the individuals.

Grid technologies constitute a successful alternative for solving those issues. This chapter presents the current status of grid technologies applied on epidemiology and how the above issues are solved in several state-of-the-art international projects.

The objective of the chapter is to make a survey of how state-of-the-art projects are solving the problems of data management, data processing and data security in epidemiology. Since there are many different use cases and approaches, different projects are considered and analysed. It does not aim at providing an exhaustive list of projects and papers, but rather to a list of relevant alternatives that can be considered as good success stories for the use of epidemiology and grids.
Related Content

Evaluation of Cognitive Function in Migraine Patients: A Study Using Event-Related Potentials
www.igi-global.com/chapter/evaluation-cognitive-function-migraine-patients/69930?camid=4v1a

Reliability and Evaluation of Health Information Online
www.igi-global.com/chapter/reliability-evaluation-health-information-online/26354?camid=4v1a

Localization of Characteristic Peaks in Cardiac Signal: A Simplified Approach
www.igi-global.com/article/localization-of-characteristic-peaks-in-cardiac-signal/136233?camid=4v1a

Bioethics
www.igi-global.com/chapter/bioethics/63393?camid=4v1a