A Web-Based GIS for Health Care Decision-Support

Jean-Baptiste Richard [a], Laurent Toubiana [b], Loïc Le Mignot [a], Mohamed Ben Said [a], Claude Mugnier [a], Christine Le Bihan–Benjamin [a], Jean Philippe Jaïs [a], Paul Landais [a]

[a] *Paris 5 University, Faculty of Medicine, Assistance Publique - Hôpitaux de Paris, EA222, Department of Biostatistics and Medical Informatics, Necker Hospital, France*

[b] *Inserm – U707, Paris, France*

Abstract

This Web-based application allows access to the epidemiology of the demand and the supply of care concerning End-Stage Renal Disease (ESRD). It is a Web-based Geographic Information System (Web-GIS), the SIGNe (Système d'Information Géographique pour la Néphrologie), designed for the Renal Epidemiology and Information Network (REIN) dedicated to ESRD. It is a visualisation and decision-support tool. This Web-GIS was coupled to a data warehouse and embedded in a n-tier architecture designed as the Multi-Source Information System (MSIS). It provides maps matching the offer of care to the demand. It is presented with insights on the design and underlying technologies. It is dedicated to professionals and to public health care decision-makers.

***Keywords**: Web-GIS ; Multi-Source Information System ; Data Warehouse ; End-Stage Renal Disease ;Public health ; Decision-making*

1 Introduction

Epidemiological changes of End-Stage Renal Disease (ESRD) during the last decade showed an increasing incidence and prevalence. However, the magnitude of this phenomenon was not precisely known in France [1]. A Renal Epidemiology and Information Network (REIN) was then built to face this poor epidemiological knowledge of ESRD [2]. A Multi-Source Information System (MSIS) was then set up [3]. This Web-based application integrates a tool dedicated to improving our knowledge of demand and supply of care for ESRD. This project involves research units (Universities Paris 5 and Grenoble, INSERM), professionals (Société de Néphrologie, Société francophone de dialyse, Société française de Transplantation), state Agencies (Agence de Biomédecine, Institut de Veille Sanitaire, Caisse Nationale d'Asssurance Maladie, Direction de l'Hospitalisation et de l'Organisation des Soins), and patient representatives.

We implemented a Geographical Information System (GIS) to support public health decision making for ESRD. SIGNe (Système d'Information Géographique pour la Néphrologie) was dedicated to dynamically visualize and analyse ESRD demand and supply of care. It was developed according to Web-GIS, Data Warehouse and Data Mining technologies, aiming to analyse ESRD epidemiology in order to improve access to care, improving public health decision-making.

2. Material and Methods

Context

Schematically, the geography of health comprises:

- the geography of disease, which covers the exploration, description and modelling of the spatio-temporal incidence and prevalence of disease, the detection and analysis of disease clusters and patterns;
- the geography of healthcare systems, which deals with the planning, management and delivery of suitable health services, after determining healthcare needs of the target community and service catchments zones.

Geographic Information Systems (GIS) appeared useful for medical research and epidemiology [4,5,6,7]. Health geography is already indispensable for public health surveillance, seems useful to identify inequities in health care delivery, and to efficiently help allocate and monitor healthcare resources. Geographical visualizations may be helpful for both studying spatial epidemiology and helping to assess the best distribution of healthcare units based on current needs.

SIGNe aims at describing spatial and temporal trends of ESRD epidemiology (incidence, prevalence, survival…), describing the offer of care, assessing the adequacy between the supply of care and the demand. It is dedicated to improving health care decision making for ESRD care.

In order to achieve these objectives, we developed a web-based interface dedicated to ESRD professionals and decision-makers at a regional and national level. This interface includes tools that allow creating dynamically both maps and charts representing current data sets regarding epidemiological and care management aspects.

Data Warehouse building

Data collection

The Multi-Source Information System (MSIS-REIN) collects all ESRD minimal patient records with their annual follow-up, the content of which was defined by a national consensus. The main variables of these records include both epidemiological information and description of the supply of care: age at first initiation of replacement therapy, type of initial nephropathy, type of dialysis, comorbidities, handicaps, remoteness from dialysis unit, characteristics of each unit of care and territorial distribution of the nephrologists. The MSIS is currently running in 8/22 French regions: Limousin, Languedoc-Roussillon, Champagne-Ardenne, Provence Alpes Côte d'Azur, Ile de France, Centre, Midi-Pyrénées and Haute-Normandie. To ensure the completeness of data collection, controls are performed in each care unit of the region and in those of the neighboring regions in order to identify all the treated patients living in the region of interest. Clinical research assistants perform quality control regionally. They assess the MSIS data compared to the patients' medical file. The size of random samples of records to be controlled is in accordance with the size of the dialysis unit.

MSIS application

The architecture of MSIS-REIN is based on a n-tier architecture (figure 1). A universal client (1st tier) connects to a dynamic Web server (2nd tier) that is in relation with several databases (3rd tier). The information system tier accesses two databases: the identification database and the production database. The identification database accessed through an identity server stores separately nominative and administrative information. The production database stores all medical information and patients' trajectories. The identity server also provides a double-entry prevention function, adapted from the sequences matching algorithm of Needleman and Wunsch [8,9]. It can recognize a patient even if an error has been introduced in its identifying items. A secure connection and the use of an identification server warrant confidentiality and security of patient information according to the French law. After completeness is achieved and quality control is performed, data are qualified as consolidated and then integrated twice a year to the data warehouse.

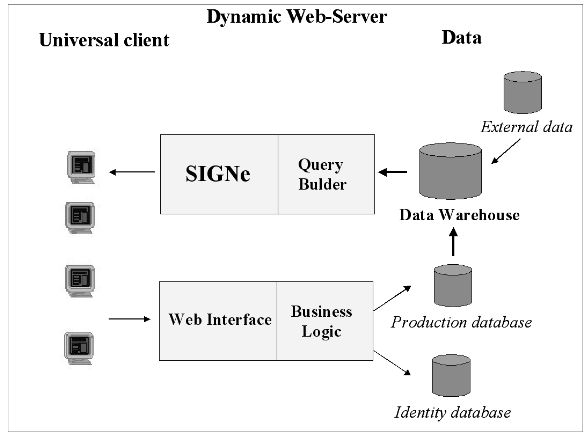

Figure 1. MSIS Architecture

Data Warehouse

The production database is designed to fit the format of a standard medical file, in order to be easily used by the physicians. However, it is not optimally organised to easily and quickly extract data. Therefore, a data warehouse has been created which builds up unified patient information, generated from distinct databases. Unlike the production database, the data warehouse is a collection of data which are subject-oriented, integrated, time variant, non volatile and organised to support analytical and decision processes [10]. Data are not created "de novo" by end-users, but derived periodically from the production database cautiously respecting privacy, integrity and confidentiality of the information.

The data warehouse integrates external data: demography, helpful in calculating incidence or prevalence rates, and geographical boundaries, in order to make spatial analysis. It is organized in several generic axes such as space, time, age of the patient, initial nephropathy, treatment method, comorbidities, transport mean…(figure 2). These axes define a n-dimensional space (or hypercube), where each item of information (fact) is mainly represented as a set of coordinates in this hyperspace. They represent the different levels on the basis of which data can be aggregated. Operations such as roll up (global view) or drill down (show details) can be performed to quickly access views at different scales. For instance, with the age axis, users can easily access data at a detailed level or at an aggregated level (e.g. classes of age). Considering the geographical level, data can be analyzed at a national, regional or health territorial level. This hierarchy allows easily accessing summarized views of our data set.

In the data warehouse, we defined two kinds of facts: selective events (e.g. patient transfer from a unit to another) and pre-processed synthesized facts, such as annual reports summarizing all characteristics of

prevalent patients. Separating these facts makes it possible to optimize queries against the database by reducing the amount of data to be scanned in the selective events table.
The data warehouse is linked to data-mining tools and to an interface providing an interactive representation of knowledge to support the decision process. Additional modules are implemented above the data warehouse: a query builder ensures transparency of the implementation and extensibility of the model; a simulation engine uses scenarios in order to build, from actual data, virtual views that will take place in the same hyperspace.

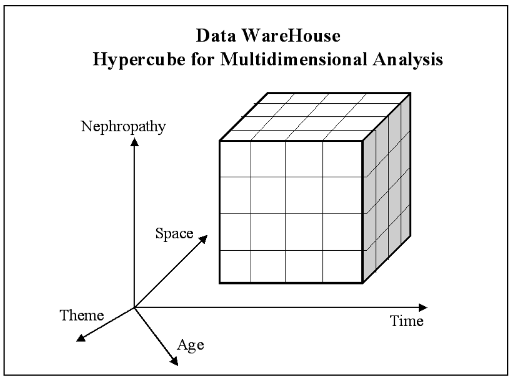

Figure 2. Data Warehouse structure

SIGNe : Data Warehouse and Web-GIS

The SIGNe project achieves a strong linkage between a geographical representation through GIS and the spatial axis of the data warehouse, in an interactive way. It uses spatial operators in order to take advantage of the spatial nature of data. After data have been geo-referenced, data sets can then be spatially selected on a map, by a single shape selection like region or health territories. Users have also a direct access to spatial aggregations, drilling or other operations on this axis.

Web-GIS

A GIS is a system composed of hardware and software used for storage, retrieval, mapping and analysis of spatially referenced (geo-referenced) information. A GIS makes possible overlaying and integrating multi-source data. It helps to discover and visualize new data patterns and relationships that otherwise would have remained invisible. It creates a link between spatial data and their related descriptive information. A Web-based GIS (Web-GIS) reproduces the main functions of GIS on a Web-interface: spatial analysis, navigation (zoom, pan), and dynamic creation of map, layer overlaying, and interactive querying. A Web-GIS is a new mean to dynamically share and represent spatial information, with a large access. A lot of Web-GIS have recently emerged (MapServer, ArcIMS, MapXtrem, Alov Map, GeoServer), using different displaying technologies: vector (SWF, SVG...), raster (PNG, JPEG...), server side or client side mapping. Regarding the interactivity, the database connection abilities and the costs of development, we chose to develop our interface with Flash MX™, PHP™ and MySQL™. Given the ease of connection to any database, this technical solution is particularly adapted to interactive cartography. This vectorial format, with a free and lightweight player, allows high levels of interactivity and a quick interface. Users can access various functionalities: zooming, selecting geographic areas, viewing additional neighboring zones by simple roll over. Moreover, this solution of a client side application is interesting considering time-response. Unlike server side applications, geographical data are loaded one time, they are easy to manipulate and do not need to be reloaded when a new request is sent.
Otherwise, we used tools to create dynamic charts in connection to the data warehouse. Different open-source programs allow the creation of these graphical representations: JPGraph, PHP/SWF Chart, or Owtchart. We used PHP/SWF Chart that creates main chart types in a vector format (SWF) compatible with our Web-GIS system.

User Interface

SIGNe interface was integrated as a new tool to the MSIS and to the data warehouse. The MSIS supplies the data warehouse with consolidated data. Once the connection to the SIGNe established, the user selects the different axes on the user interface (figure 3). For instance, the user can choose to study ESRD incidence in 2003 for one specific region of interest.

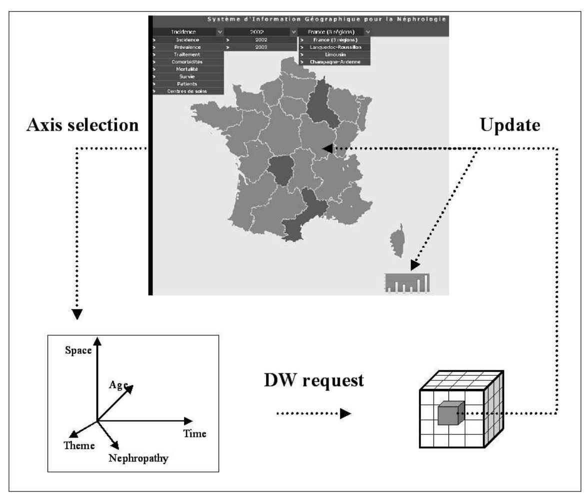

Figure 3. SIGNe user interface

Once the axis selected, a query builder transform the user request into a standardized SQL syntax, which allow the correct data from the Data Warehouse to be extracted. The user also chooses either to construct a map or to graphically represent the current data set.

3. Results

Data set

The data collection began in 2002 with three French regions. Today, the MSIS is applied in eight French regions. More than 400 health care professionals from 174 dialysis units participate in data collection. The evolution of the number of ESRD patients in the production database is presented on figure 4. Currently, the data set gathers more than 13,000 records, that represent around one third of the dialyzed ESRD French population.

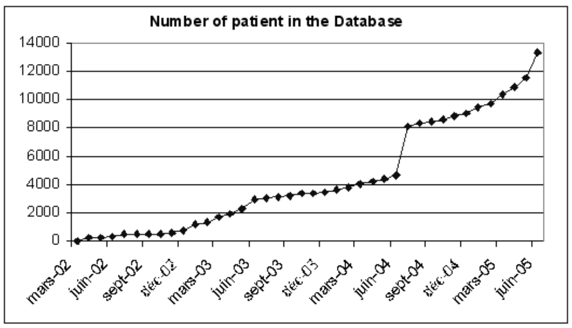

Figure 4. Evolution of the number of patients' records

Epidemiological results

The thematic map (figure 5) is a dynamical visualization where the attributes of geographic features (here the incidence rates by region) are displayed on a map. It allows assessing the geographical distribution of new ESRD cases considering several French departments. Other thematic maps can be obtained for different themes (prevalence, mortality...), and also at a smaller scale. For instance, the user is able to map the disease distribution by district or even smaller administrative boundaries.

The pie chart (figure 6) is aimed at defining more precisely the ESRD population. It represents the distribution of incidence by age group. Other similar representations can be created regarding the type of treatment, the co-morbidities or the handicaps.

Health care attractiveness

Figure 7 shows an example of both supply and demand of care representation. Patients' affiliations according to their district of residence (dots) are linked to the location of their dialysis unit. It shows for example, that the hospital in Reims attracts patients beyond the limits of the Marne department despite the presence of other dialysis units in the adjoining departments.

Such a map provides medical professionals and administrators with informations about matching the supply and demand of care for dialysis. Patients are geo-referenced using their place of residence and the location of their dialysis unit.

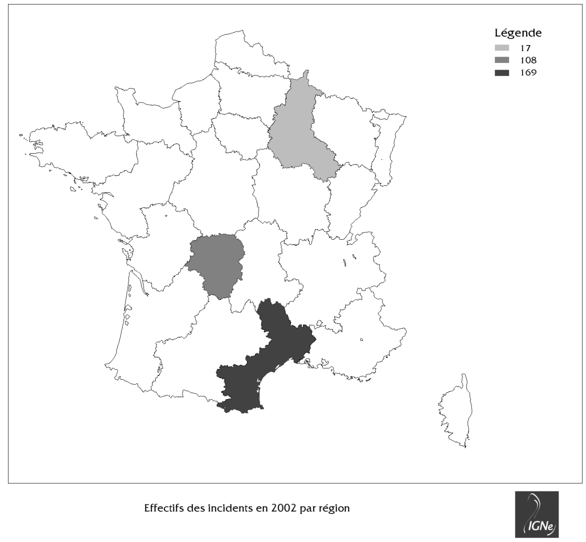

Figure 5. Cartographic representation

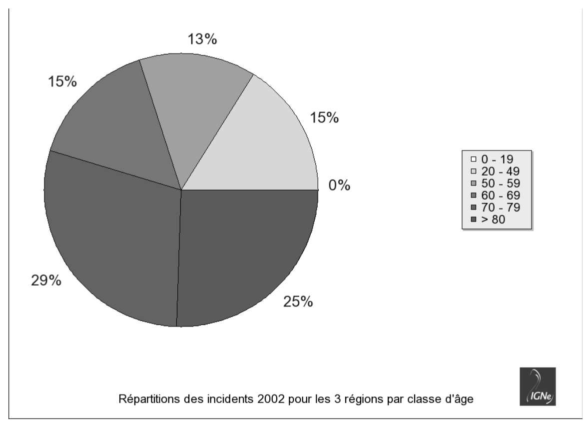

Figure 6. Pie Chart representations

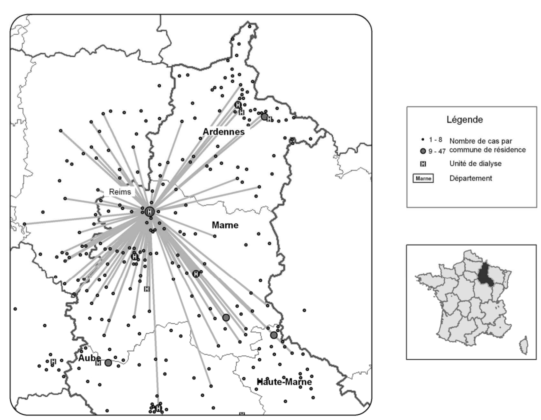

Figure 7. The Reims hospital catchment zone

4. Discussion

Here we present results concerning the epidemiology of the demand and the supply of care for End-Stage Renal Disease (ESRD) and the catchment areas of healthcare facilities as analysed by the SIGNe. Considering currently the restricted number of data available at a temporal level, we were unable to develop a temporal analysis. However, these kinds of results will be very interesting in analysing incidence and prevalence trends and temporal queries have already been prepared. Moreover, in order to better study the mismatch of the supply to the demand, we are currently working on modelling patient travel times to reach their dialysis unit, with two main objectives. The first one is to represent access to healthcare considering both the patients' place of residence and also their travel time. The second one is to propose regional scenarios for supporting new distributions of the offer of care according to the geography of the demand.

The use of web-based interfaces including interactive mapping is growing in health geography. Such interfaces are usually developed in order to help public health surveillance and a better knowledge of disease geographical distribution, such as the application of the world health organization [12]. Such applications were seldom developed as a support of health care decision-making.

5. Conclusion

SIGNe offers a dynamic interface for accessing and contributing to health care information concerning ESRD. It allows the representation of the demand as well as the supply of care. Moreover, it helps describe the current match between the location of care and the place of residence of ESRD patients. It is thus a support to health care decision-making.

We are currently working at describing more precisely accessibility to health care units, in order to support health care planning.

6. Acknowledgments

We thank the members of the participating regions to the MSIS. This research was funded by a grant from STIC-Santé-Inserm 2002, n°A02126ds, and by Paris 5 University. This work was also supported by grants provided by the Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, the Institut de Veille Sanitaire and the Etablissement français des Greffes. Xavier Ferreira and Jean-Philippe Necker are acknowledged for their skilful help.

7. Reference

[1] Landais P. End-Stage Renal Disease in France: Epidemiology of care demands. Supply of care and prevention. Presse Med 2002; 31:167-85.

[2] Landais P, Simonet A, Guillon D, Jacquelinet C, Ben Saïd M, Mugnier C, Simonet M. SIMS@REIN: a multi-source information system for end-stage renal disease. CR Biol 2002; 325: 515-528.

[3] Ben Said M, Simonet A, Guillon D, Jacquelinet C, Gaspoz F, Dufour E, Mugnier C, Jais JP, Simonet M, Landais P. A dynamic Web application within an n-tier architecture: a Multi-Source Information System for end-stage renal disease. Stud Health Technol Inform 2003;95:95-100.

[4] Croner CM. Public Health GIS and the Internet. Annu Rev Public Health 2003; 24:57-82.

[5] Cromley EK. GIS and Disease. Annu Rev Publ Health 2003; 24:7-24

[6] Richards T.B. et al. Geographic information systems and public health: mapping the future. Public Health Rep 1999; 114: 359-60.

[7] Bedard Y, Gosselin P, Rivest S, Proulx MJ et al. Integrating GIS components with knowledge discovery technology for environmental health decision support. Int J Med Inform 2003;70:79-94.

[8] Needleman S, Wunsch C. A general Method Applicable to the Search for Similarities in the Amino Acid Sequences of Proteins. J Mol Biol 1970: 48;443-453.

[9] Mugnier C, Ben Said M, Le Mignot L, Jais J, Le Bihan C, Landais P. Avoiding doubles in distributed nominative medical databases: optimization of the Needleman and Wunsch algorithm. MEDINFO 2004 2004(CD):1763

[10] Jones K. An Introduction to Data Warehousing: What Are the Implications for the Network? Int J Network Mgmt 1998;8:42-56.

[11] http://www.usrds.org/odr/xrender_home.asp (last access : 01/07/2005)

[12] http://www.who.int/csr/mapping/en/

(last access : 01/07/2005)

8. Address for correspondence

Paul LANDAIS, MD PhD, Service de Biostatistique et d'Informatique Médicale, Hôpital Necker-Enfants Malades. 149, rue de Sèvres, 75743 Paris cedex 15. e-mail : landais@necker.fr