Decision support system for health care resources allocation

Abderrazak Sebaa[1], Amina Nouicer[2], AbdelKamel Tari[1], Ramtani Tarik[2], Ouhab Abdellah[2]

[1] LIMED Laboratory, Computer Science Department, Bejaia University, Algeria
[2] Computer Science Department, Bejaia University, Algeria

Type of article: Original

Abstract

Background: A study about healthcare resources can improve decisions regarding the allotment and mobilization of medical resources and to better guide future investment in the health sector.

Aim: The aim of this work was to design and implement a decision support system to improve medical resources allocation of Bejaia region.

Methods: To achieve the retrospective cohort study, we integrated existing clinical databases from different Bejaia department health sector institutions (an Algerian department) to collect information about patients from January 2015 through December 2015. Data integration was performed in a data warehouse using the multi-dimensional model and OLAP cube. During implementation, we used Microsoft SQL server 2012 and Microsoft Excel 2010.

Results: A medical decision support platform was introduced, and was implemented during the planning stages allowing the management of different medical orientations, it provides better apportionment and allotment of medical resources, and ensures that the allocation of health care resources has optimal effects on improving health.

Conclusion: In this study, we designed and implemented a decision support system which would improve health care in Bejaia department to especially assist in the selection of the optimum location of health center and hospital, the specialty of the health center, the medical equipment and the medical staff.

Keywords: Data warehouse, Health sector, Decision support system

1. Introduction

1.1. Background and the problem specification

Emerging countries, especially the least-developed countries are severely lacking in health infrastructure and medical staff. These countries also suffer from hospitals that are under-equipped. The evidence of health inequalities between the poor and non-poor countries and population are detailed in (1). Decision-making about the alternative uses of health care resources is an issue of critical concern for governments and administrators in all health care systems (2). This situation urges governors and policy makers to find solutions. Algeria is one of these emerging countries and it endeavors to find a solution to this same problem. Therefore, the Algerian director health scheme plans to invest $20 billion between the periods 2009 and 2025 (3). The investment plan includes construction and modernization of health centers and hospitals, acquisition of new medical equipment, training and recruitment of health body. The detailed plan of this Algerian health scheme plans to achieve 172 general medical centers 45 specialized medical centers, 377 poly-clinics, 1000 regional health centers, 17 paramedical training schools, and more than 70 disability institutions. However, the issue was how to achieve this plan with a maximum equality between its people. A medical decision support system is a health information technology system which assists and provides doctors and other health professionals with clinical decision support. Using data warehousing technology, it is one of the most common ways to achieve a decision support system. A data warehouse (DW) is a system used

for reporting data analysis, and is considered as a core component of business intelligence (4). It is a collection of data, organized to be used as a decision support (5), such as data characterized by these factors:

- Data are organized by themes or subjects, gathering all information related to a specific topic, facilitating decision making.
- Data are mainly used in consultation mode and are less frequently updated or deleted by users. This keeps traceability of information, performing analysis over a long time.
- Integration of data (6) eliminates all conflicts of representation; syntactic and semantic, to get a uniform and consistent representation of data when data are loaded at the Data Warehouse.

1.2. Objectives

The focus of this study was to achieve a decision support system which performs apportionment and allotment of medical resources of Bejaia Health Sector. In fact, Bejaia health care institution is in constant growth. Supervising and monitoring such evolution requires an effective solution. This study detailed the design and implementation of the Data Warehouse, which is used to provide information to analysts and policy makers. The aim of designing and building this warehouse was to lead to a platform for performing fair apportionment and allotment of medical resources, by providing the decision-makers a clear view of the health sector data in the Bejaia department. It will help decision-makers to make better decisions regarding the location of health centers and their specialties, medical staff recruitment plans, most required equipment for hospitals and improving the efficiency of medication delivery.

2. Material and Methods

This research study was carried out at Bejaia health sector institutions. Indeed, we integrated existing clinical databases from different Bejaia department health sector institutions (an Algerian department) to collect information about patients from January 2015 through December 2015. The used materials and methods detailed in this section.

2.1. Materials

We made the choice to use a data warehousing tool technology based on the relational database. The designing of the database was carried out with the multi-dimensional model. The implementation of the data warehouse was performed using SQL server. Reporting tools were performed with MS Excel and a set of Microsoft Visual Basic (VBA) programs designed for automatic querying of the database.

2.2. Methods

Designing and implementation of a data warehouse was a complex task which consisted of several steps: requirements identification, data sources analysis and extraction, Integration, loading at the warehouse and establishing various reporting tools (analysis, data mining or queries). In this project, two kinds of requirements were identified: functional requirements which were meant to look for decisional information about diseases and patients and health centers, such as patient information, the name and type of diseases, disease patterns across time, or in location, and equipment or ward, concerned. Non-functional requirements which were meant to ensure the processing speed, performance and security.

2.2.1. Data sources and Study population

Data extraction was to retrieve data from several heterogeneous and distributed sources. It required the synchronization of extraction processes to ensure data integrity loaded. These sources may have included databases, data files, external sources, etc. Before being stored, data sources had to first be cleaned. The cleaning process was to select and purify the data to eliminate problems and reconcile the semantic differences between these data. In our case, the data were coming from several sources of Bejaia hospital information systems which stored data about patients of Bejaia department. In this study, we limited data collection from January 2015 through December 2015. Bejaia is one of the 48 Algerian departments (Bejaia health scheme is a part of the Algerian director health scheme). It has a population of around 1 million people, it covers 51 municipalities. Bejaia health sector employs 33 university hospital staff, 245 specialist medical practitioners, 734 general medical practitioners, and 2,742 paramedical staff. Its mission is to support by an integrated and hierarchical manner, the health needs of the population. It has the following institutions: A university hospital, 15 public hospital institutions, a specialized public hospital institution, 51 polyclinics, 8 proximity public institutions of health, several laboratories, a blood transfusion center, 30 care units nearby, and a paramedical school. Each new patient, doctor, or disease must be loaded into the HIS (hospital information system) of Bejaia. So, all information about patients, doctors their addresses, diseases, visit date and the visited health center must be stored in the tables shown in Figure 1 which are:

"Patient", "Doctor", "Disease", "Period", "Health Ward" and "Location". The next step was the data model, which was required to meet planned needs and be easily expandable to meet future needs.

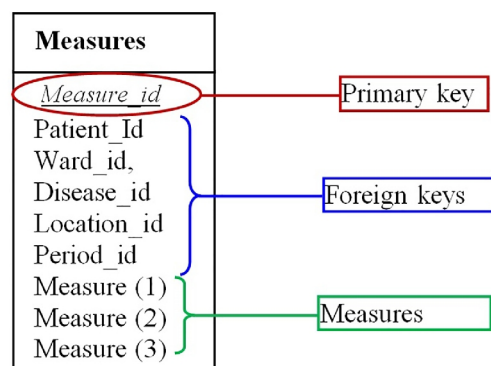| Patient	Period	Health-ward
patient_id patient_sex, patient_ age patient_ address,	*Period_id* Period _duration Period _beg_day Period _beg_month Period _beg_Year	*Ward_id* Hospital_Id Ward_specialty Ward_capacity Equipment quality
Disease	**Doctor**	**Location**
Disease_id Disease_name Disease_ type Disease_symptom	*Doctor_id* Recruitment_dat Degree Specialty	*Location_id* Street City Town

Figure 1. Dimension tables

2.2.2. Data preparation modeling
Entity-Relationship diagrams and normalization techniques are popularly utilized for database design for data warehouses. However, the database designs recommended by Entity-Relationship diagrams were inappropriate for decision support systems where efficiency in querying and in loading data was important (7). Thus, we chose to use the multi-dimensional model. To create the multi-dimensional model, we started with the creation of the warehouse database by creating and populating dimension and fact tables.

2.2.3. Data and schema structure
We used a star schema to represent the multidimensional data model. Indeed, in the star schema, the database consisted of a single fact table and a single table for each dimension. Each tuple in the fact table consisted of a pointer (foreign key - often uses a generated key for efficiency) to each of the dimensions that provided its multidimensional, coordinates, and stored the numeric measures for those coordinates. Each dimension table consisted of columns that corresponded to attributes of the dimension (8). Materialized views were chosen to store the integrated data of dimensional tables. In fact, a materialized view is a database object that contains the results of a query, or a duplication of one or more tables. Generally, we used materialized view since it allowed being modified in different ways: (to change its storage characteristics, to change its refresh mode or time. to alter its structure so that it was a different type of materialized view). Indeed, in our case, some information about patients and health staffs could change over time. For example, the patients address. Materialized view maintenance is to keep materialized view up to date when the underlying data sources are updated over time (7, 9). Several maintenance methods were proposed in the literature. Therefore, using a view maintenance method could eliminate the data change issue.

Measures	
Measure_id	Primary key
Patient_Id Ward_id, Disease_id Location_id Period_id	Foreign keys
Measure (1) Measure (2) Measure (3)	Measures

Figure 2. Fact table.

As illustrated in Figure 2, fact table (Measures) contained mainly three information parts: the primary key "Measure_id", foreign keys, Patient_id, Doctor_id, Disease_id, Location_id, Period_id, Ward_id" and a set of measures such as the number of patients in a given location and period. The set of tables presented previously were used to create the star schema shown in Figure 3. The center of the star is the fact table "Measures" and the points of the edge of the star are the dimension tables "Patient, Doctor, Disease, Period, location, and health-ward". The denormalized structure of the dimensional tables in star schemas is more appropriate for browsing the dimensions.
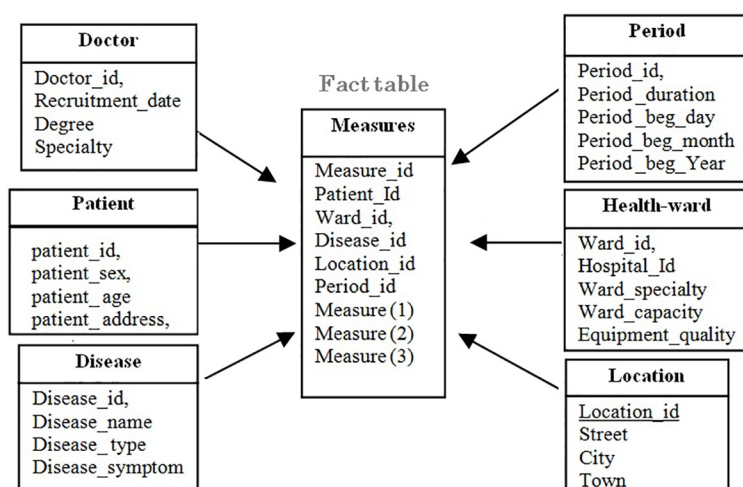


Figure 3. The star schema.

2.2.4. Data transformation

Data transformation is a set of rules to format and convert extracted data from one format (e.g., database file, XML document, or Excel sheet) according to the target warehouse schema, such as assigning semantics to data sources and combining source fields to target fields. So, the transformation was to map the incoming data from different sources to be in a format of Bejaia data warehouse. We used some techniques, such as combining multiple name fields into one field or breaking down fields, mapping creation which was to determine the relationship between data elements of two applications. We updated some column names into other names and convert columns from one data type to another.

2.2.5. Data loading and OLAP Cube building and deploying

Data loading was the last step of ETL process. It consisted of loading data into the data warehouse for analytical purposes (10). Information about creation, management and use of the warehouse, about sources and their contents, the warehouse schema, refresh rules, profiles and user groups (11) were also stored in a separate directory. This information was called "Metadata". The data warehouse was dedicated to decision analysis OLAP-style. It provides access to the warehouse, converts client requests for access to the DW and provides multidimensional views of data to support tools decision (12). The multi-dimensional model supported OLAP design to provide instant results of analytic queries.
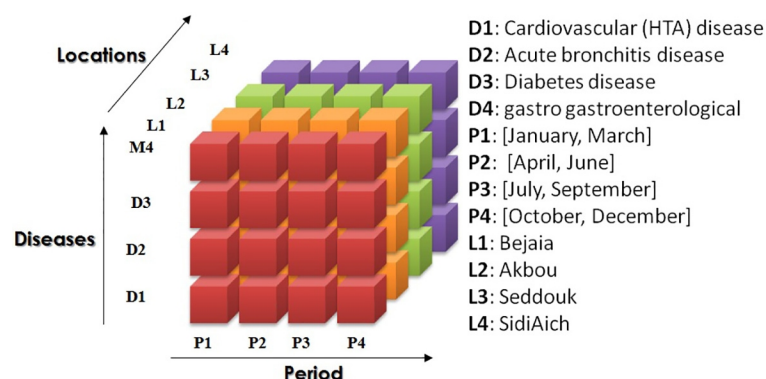
Figure 4. Data cube.

To facilitate data analysis, OLAP transformed the medical data stored in the relational database to relevant medical information and made it easy to operate, through creating a data cube. It was the objective of this step, i.e. the design of our multidimensional database for OLAP cubes. Therefore, to design our warehouse we used MOLAP (Multidimensional OLAP) together with T-SQL language. Since we had six dimensions and each dimension had at least four attributes, there may have been several cubes depending on the chosen dimensions. A data cube was identified by three dimensions or an attribute of a dimension. Figure 4 shows a representation of a data cube by using the dimensions: Disease (D1, D2, D3, and D4), Period (P1, P2, P3, and P4), and Location (L1, L2, L3, and L4).

2.2.6. Data Warehouse implementation
Based on several parameters such as: administration tools, protection of brute force, deployment and maintenance, data confidentiality and integrity, concurrency control, GUI, data backup and restore, availability and mastery of the software, we compared several data warehouse building tools: Oracle, Pentaho, Talend Open Studio and SQL Server. The latter was chosen to build the data warehouse, and using Microsoft Excel, we created reporting tools.

2.2.7. Data analysis
The aim of our medical decision support system was to address several problems which covered several aspects: public health, financial, economic and management of the health sector in Bejaia department. It allowed us to give important orientations during the public health planning stages, to have a complete predictive vision by a better repartition of care offerings (hospital specializations, apportionment and allotment of medical equipment , number of specialist per hospital, number of beds by location, number of beds by specialty) these offers need to routinely be adjusted with changing of interest (changing wellbeing procedures, sicknesses, wellbeing structures, populace age and address of the populace). Our data warehouse will help decision makers toward settling on target choices which will prompt a huge decrease in cost, more productive strategy for precise wellbeing administration as quickly as time permits, enhance the accessibility of material and HR, improve the services quality, and patients' safety.

2.3. Ethical aspects
Protection of personal information about patients: Patients first name, last name, place, and date of birth were removed from our system. They were replaced by an identification and age of the patient. Only those who were involved in the direct system development and provision of care and the medical managers had access to items of medical information. Data transfer to the data warehouse was designed to occur after the suppression of the individual identity information of patients.

3. Results
At that point, we achieved a medical data warehouse on which we accumulated medical information about Bejaia Health sector, and a reporting tool based on a graphical presentation of information. By ensuring a connection between our data cube and the reporting tool, health professionals and decision-makers can use the decision support system easily and in real time. The principle interface of the reporting tool appeared in Figure 5, with which, users can get other GUIs which provide information on various health parameters.

Figure 5. Main reporting interface.

Users can obtain reports between several medical parameters which will be analyzed. So, the system will give assistance and reports about, for example, the optimal place to build or to extend health centers in a given specialty. It was possible to analyze the collected data and to create a histogram of each disease with the different cities. In Figure 6, a histogram represents each disease in a specific city. Using the information collected in this figure, we can improve the distribution of medical specialists in different cities.

Figure 6. Reporting about distribution of diseases in different cities.

4. Discussion

Although several studies (13, 14) have suggested the use of medical data to identify health care trends, prevent diseases and combat social inequality, little work has addressed the apportionment and allotment of medical resources issues using data warehousing technology. This is due, firstly, to the slow adoption of data warehousing technology in the medical sector (15). Secondly, is that most of the studies on medical data warehousing are directed towards specific diseases. Indeed, epidemiological data warehousing was presented by Kerkri et al (16) based on EPIDWARE architecture. They contended that EPIDWARE design will facilitate statistical and epidemiological studies. A hepatic transcriptome data warehousing was presented in (17). It was a part of a bioinformatics project called GADAW. GADAW team collected within a single knowledge base, complex, varied and numerous data of the liver genes for analysis. Another example is a data warehouse for diabetes treatment which was proposed in (18). In Rwanda, a data warehouse (19) to the Rwandan heath sector based on Electronic Health Record was proposed for a swift decision-making process through discovery, then alerting about the epidemic issues. In Austria, DWs of different health insurance organizations were merged in an evidence-based medicine collaboration project (20), called HEWAF (Healthcare Warehouse Federation). The French health policy developed the PMSI (21) project (Program for a Medical Information System) which was aimed at the development of cost analysis based on diagnosis related groups. Since inequalities of health resources distribution mainly affect poor countries, the apportionment and allotment of medical resources using a decision support system have received little interest. This led us to implement the first solution (22) on which a data warehouse for optimal allotment of medical resources was proposed. However, only data coming from health institutions of Bejaia city were considered. So, this solution disregarded the ability of patients to move between the cities of Bejaia department, which reduced the accuracy of the decisions proposed by the decision support system (22) for the distribution of health resources. This work took into account all health centers of Bejaia department as data sources on which we collected data. The proposed solution is able to handle varied data sources, diminish the dimensionality constraint, and handling of multiple information factors. However, the best decisions to address inequalities of health resources distribution can be taken when we extend this work at the national level.

5. Conclusions

The medical decision support system is one of the most promising fields, and new technologies are being designed for its easy implementation. In this paper, we described an approach to conceptual warehouse design which manages and coordinates medical resources. The proposed decision support system gives a lot of information on epidemiology and public health, about various diseases, their concentration, and repartition in Bejaia department. It shows age classes, periods, and areas affected by a specific disease etc. We have provided an interesting solution at many levels (economic, social, technical, political...), it will be the support for data analyses for decision taking for health sector development in Algeria and especially in the Bejaia region. However, extending this solution at the

national level will give better results. Moreover, due to the sensitive nature of medical data, it is important to take into account their own privacy policies problems.

Conflict of Interest:
There is no conflict of interest to be declared. This article is among the best papers of the International Conference on Health Sciences and Medical Technologies, Tlemcen (Algeria), 2016. The article is published free of charges.

Authors' contributions:
All authors contributed to this project and article equally. All authors read and approved the final manuscript.

References:
1) Wagstaff A. Poverty and health sector inequalities. Bull World Health Organ. 2002; 80(2): 97-105. PMID: 11953787, PMCID: PMC2567730.
2) Ross J. The use of economic evaluation in health care: Australian decision makers' perceptions. Health Policy. 1995; 31(2): 103-10. doi: 10.1016/0168-8510(94)00671-7, PMID: 10141251.
3) ANDI: National Agency for Investment Development of Algeria. Available from: http://www.andi.dz/index.php/fr/secteur-de-sante.
4) Dedić N, Stanier C. An Evaluation of the Challenges of Multilingualism in Data Warehouse Development. 18th International Conference on Enterprise Information Systems. 2016; 1: 196-206. doi: 10.5220/0005858401960206.
5) Inmon WH. Building the Data Warehouse. 2nd edition. New York: John Wiley and Sons; 1996.
6) Shah SP, Huang Y, Xu T, Yuen MM, Ling J, Ouellette BF. Atlas–a data warehouse for integrative bioinformatics. BMC bioinformatics. 2005; 6: 34. doi: 10.1186/1471-2105-6-34. PMID: 15723693, PMCID: PMC554782.
7) Zhang X, Ding L, Elke A, Rundensteiner EA. Parallel multisource view maintenance. VLDB. 2004; 13(1): 22-48. doi: 10.1007/s00778-003-0086-0.
8) Surajit CH, Dayal U. An overview of data warehousing and OLAP technology. ACM. 1997; 26(1): 65-74. doi: 10.1145/248603.248616.
9) Sebaa A, Tari A. Self-Maintainable Views Group in P2P Environment. American Academic & Scholarly Research Journal. 2014; 6(4): 186.
10) Golfarelli M. From User Requirements to Conceptual Design in Data Warehouse Design – a Survey. In: Data Warehouse Design and Advance Engineering Applications: Methods for Complex Construction. IRMA International. 2010; 1-16. doi: 10.4018/978-1-60566-756-0.ch001.
11) Peter R, Coronel C, Morris S. Database systems: design, implementation and management. 10th ed. Cengage Learning: Wadsworth Publ; 2013.
12) Lightstone SS, Teorey TJ, Nadeau T. Physical Database Design: the database professional's guide to exploiting indexes, views, storage, and more. Morgan Kaufmann. 2010. ISBN:13-9-78-0-12-369-389-1.
13) Ewen EF, Medsker CE, Dusterhoft LE. Data warehousing in an integrated health system: building the business case. ACM. 1998: 47-53. doi: 10.1145/294260.294271.
14) Kuo MH, Sahama T, Kushniruk AW, Borycki EM, Grunwell DK. Health big data analytics: current perspectives, challenges and potential solutions. International Journal of Big Data Intelligence. 2014; 1(1-2): 114-26. doi: 10.1504/IJBDI.2014.063835.
15) Sumathi S, Esakkirajan S. Fundamentals of relational database management systems. Studies in Computational Intelligence. 2007. ISBN: 978-3-540-48397-7.
16) Kerkri EM, Quantin C, Allaert FA, Cottin Y, Charve P, Jouanot F, et al. An approach for integrating heterogeneous information sources in a medical data warehouse. J Med Syst. 2001; 25(3): 167-76. doi: 10.1023/A:1010728915998. PMID: 11433546.
17) Guérin E, Marquet G, Burgun A, Loréal O, Berti-Equille L, Leser U, et al. Integrating and warehousing liver gene expression data and related biomedical resources in GEDAW. In: International Workshop on Data Integration in the Life Sciences. 2005: 158-74. doi: 10.1007/11530084-14.

18) Pedersen TB, Jensen CS. Research issues in clinical data warehousing. In: Scientific and Statistical Database Management. IEEE. 1998: 43-52. doi: 10.1109/SSDM.1998.688110.

19) Pavalam SM, Jawahar M, Akorli FK. Data warehouse based Architecture for Electronic Health Records for Rwanda. In: Education and Management Technology (ICEMT). IEEE. 2010: 253-5. doi: 10.1109/ICEMT.2010.5657660.

20) Banek M, Tjoa AM, Stolba N. Integrating different grain levels in a medical data warehouse federation. In : International Conference on Data Warehousing and Knowledge Discovery. Springer-Berlin, Heidelberg. 2006: 185-94. doi: 10.1007/11823728_18.

21) Olive F, Gomez F, Schott AM, Remontet L, Bossard N, Mitton N. [Critical analysis of French DRG based information system (PMSI) databases for the epidemiology of cancer: a longitudinal approach becomes possible]. Rev Epidemiol Sante Publique. 2011; 59(1): 53-8. doi: 10.1016/j.respe.2010.09.001. PMID: 21251779.

22) Sebaa A, Tari A, Ramtani T, Ouhab A. DW RHSB: A Framework for Optimal Allocation of Health Care are Resources. International Journal of Computer Science, Communication & Information Technology. 2015; 2(1): 12-7.