Optimizing Disease Management with Data Warehousing

Azadeh Nazeri[1](MD), Mahtab Karami[2], Reza Safdari [3](PhD),Majid Yaghoubi Ashrafi[4](MD)*

[1] MD, Mazandaran University of Medical Sciences, Sari, Iran
[2]PhD student of Health Information Management, school of Allied-Medical sciences, Tehran University of Medical Sciences, Tehran, Iran
[3] PhD of Health Information Management, Associate-professor, school of Allied- Medical sciences, Tehran University of Medical Sciences, Tehran, Iran
[4]*MD, Vice-chancellor for treatment, Mazandaran University of Medical Sciences, Sari, Iran.
* Corresponding Author: Email: beh.yaq@gmail.com

Abstract: Diseases management offers the benefits of lower disease occurrence, improved patient care, and lower healthcare costs. The key mechanism used to identify individuals at risk, stratify patients by risk level and track patients' progress through the disease management is the data warehouse (DW). By data warehouse, healthcare organizations, clinicians and managers can reduce costs generated by process redesign, obtain right information in areas of patient care, budgeting, planning, research, process improvement, external reporting, benchmarking, trend analysis and marketing and make the right decision to prevent the progression or appearance of the illness, thus maintaining the health of the population. This review article describes operational systems, DW features and component and role of it in disease management programs.
.[Azadeh Nazeri, Mahtab Karami, Reza Safdari,Majid Yaghoubi Ashrafi**, Optimizing Disease Management with Data Warehousing.**. *Life Sci J* 2013; 10(4): 929-932](ISSN:1097-8135). http://www.lifesciencesite.com.118

Key words: disease management; data warehouse; OLAP; OLTP; data mining; clinical decision making

Introduction

Disease management (DM) is a set of tasks which are centered on the data obtained from the health risk appraisals, general health assessments and satisfaction of patient and clinician. By integrating these data can determine patient's risk, treatment and health status. This means that risk; treatment and health statuses focus on early identification of patient at risk, rate of compliance with treatment protocols and severity of illness or patient's episodes of care, respectively [1]. Therefore the aim of DM programs is to identify patients or susceptible to disease and implement the programs to prevent disease progression or appearance which lead to public health [2].The Successful DM program depends on the correct application of "information technology". In this regard, the data warehouse (DW) as a decision-making tool plays important role and it is a key mechanism to detect people at risk, stratify patients by risk level and monitor the disease progression in DM. the proper application of DW comprises some benefits for both individuals participating in DM and healthcare organization such as cost reducing and quality increasing in patient's life [1-3]. Since DW is a system that is designed and built separates from the clinical operation systems but is fed from such systems [3-4]. The aim of this article is to describe DW and its effects on optimizing the DM.

Operational system Vs. Data warehouse

The operational systems (OS) are online transaction processing systems that are used to perform basic daily activities in organizations. Each transaction is processing information about an entity such as a physician order, a billing or a disease [2-5].These systems have been produced to provide strategic information and to get this kind of information should only be used Decision Support Systems (DSS) with specific tools. Since the beginning of 1990 in the field of DSS, Technologies like DW and online analytical processing (OLAP) developed to facilitate the analysis in DSS [6-8]. Bill Inmon- the father of data Warehousing- defines DW as "a subject-oriented, integrated, nonvolatile, and time variant collection of data in support of management's decisions." Based on this definition, the data in DW have key features as follow:

-Subject-oriented Data: against OS which store data by individual applications, in DW, the data are stored by subjects, not by applications. These subjects include patient, physician and etc [6, 9-12].

-Integrated Data: it means all the data related to a subject comes from several OS. Then the inconsistencies are removed from them. Finally the relevant and cleanings data are composed, stored and analyzed. For example all of patient data such as clinical, financial, administration, medical images and laboratory results are integrated and stored by patient subject [6, 9-12].

-Nonvolatile Data: in DW, the data can not update in real time. The data are moved from OS into DW at specific intervals depending on the requirements of organization. These data are read only and the users can't manipulate them such as delete, add or change. For example the data which are relevant to different episodes of treatment cannot be changed [3,6,9-12].

-Time- variant Data: in primarily, the data in DW is meant for decision making and analysis. Because of this purpose, the data have to store historically. This feature enables the users to perform past analysis, relate information to the present and forecast the future. For example, in DW, there are the history of patient hospitalization and the times which a patient has participated in DM programs. It is provide the clinician to comprise the various care episodes with each other [6,9-12].

-Data Granularity: data granularity in DW refers to the particular level of data details and values. For example, in DM, the patient responses can be observed in different hours, minutes and seconds.

Also, the DW is collection of several data mart. The data mart is a subject-oriented database which is relevant to one or group of processes related to specific activities. On the other hand, the data mart is department-oriented. For example a clinical DW can be combined of various data marts such as radiology, laboratory and pharmacy [12-13].

The Data warehouse Architecture

Ponniah defines architecture as "The structure that brings all the components of a data warehouse together is known as the architecture. The architecture defines the standards, measurements, general design, and support techniques." DW architecture groups components into the three areas including data acquisition, data storage, and information delivery [6]. Building these areas require the components that are necessary to arrange properly in order to meet the goals of organization. These components are described in the following

1-Source Data: these data fall into four categories. The produced data which are came from various OS of organization. The internal data includes private documents and customer profiles, etc. The archived data which are in different format such as diskette, tape cartridge, microfilm or CD. The external data refers to statistics published by external agencies, standard values of financial and quality indicators to check on organization performance [6,12-15].

2-Data staging: in this step three major function need to be performed as follows:

***-Data extraction*:** this function has to deal with several data sources and diverse data formats. That means the source data may be in relational database, legacy network, hierarchical models or flat files. To perform extract function in the separate environment need to be used proper techniques to move the data into DW.

-Data cleansing: this function includes correction of misspelling and conflicts among data, elimination of duplicates, and providing default values for missing data elements.

-Data Standardization: this function forms a large part of data transformation. Standardization of data is performed in two forms. First, the standardization of the data types and field length for the same data elements form multiple data sources. And second, the semantic standardization. The later means to resolve the problems related to synonyms and homonyms [9, 12-15].

-Data summarization: this function involves purging source data that is not useful, sorting and merging of data in the data staging area. When the data transformation function ends, a collection of integrated, cleaned, standardized, and summarized data is ready to load into DW [9, 12-15].

3- Data storage: the data storage for DW is separate from OS, and must be in structures suitable for analysis. The data structure on DW is depending on how it is used for complex analysis. For example, if data used for research and cub functions, the OLAP structure is appropriate. But if the data used for trend analysis, observation of standard reports and data screening it would better use OLTP structure. OLTP structure is relational tables in relational databases. It means two dimensional format, row and column. On the other side, in OLAP structure, the data are multidimensional that known as cube [5, 17-18]. You can see cube in figure (1).

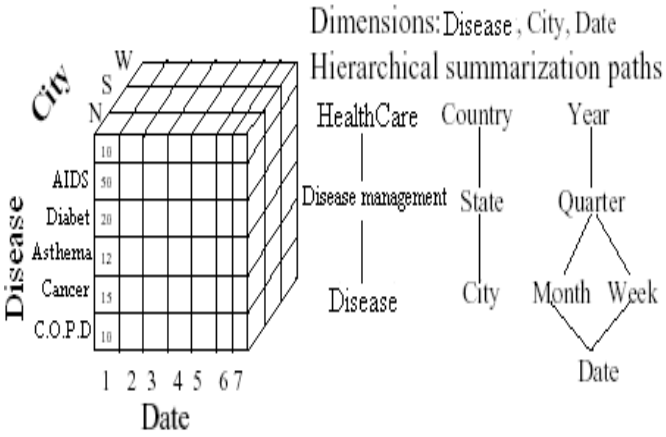

Figure1. The cube structure in DM

The OLAP structure has STAR schema. It is composed of the fact table in the middle surrounded by the dimension tables. The measurements, called facts are those numbers of interest to the analyst. For business these numbers are usually dollar amounts

and quantity counts. These numbers can be added together in meaningful ways analytically.

There are many numeric values that represent the facts. However, many of the facts that were identified were not meaningful when added together as they would be in a typical business case. In DW we call such facts non-additive. The other type of facts Factless facts are not measurements, but rather record the occurrence of an event. In this case, the events were hospitalization and emergency room visits. Dimensions are what give meaning to the numbers and measures and are the ways in which the facts are viewed and analyzed [5, 18-19]. Fact table and dimensions are shown in figure (2).

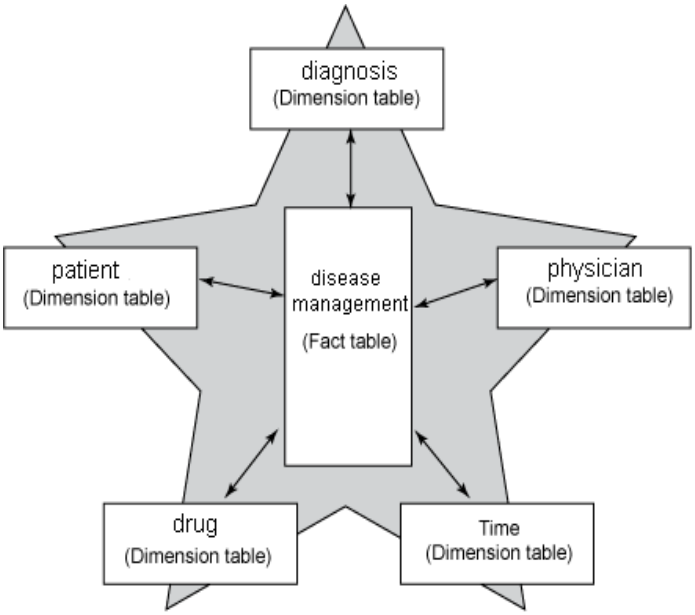

Figure 2: fact and dimensional table

4- **Information delivery**: the DW includes different methods of information delivery using internet and intranet such as multidimensional analysis, data mining, ad hoc reports, etc. the users of DW includes analyst, clinician and manager can be able to provide query and report online, and receive regularly reports via email [6,10,15,21].

5- **Metadata:** Metadata is data about data. The information about the data elements such as name, size, data type, field length, place, as well as security levels in DW [6, 20, 22].

- Operational metadata contain all of information about operational data sources.
- Extraction and transformation metadata contain all data about data extraction and transformation such as the extraction frequencies, methods, business rules, etc.
- End-user metadata is the navigation map of DW. It enables the users to look for information using their own terminology [6].

6- **Data management and control:** This component of the DW architecture is very important and coordinates the services and activities within the DW such as data extraction, data transformation, and data quality and data security. As well as the management and control component interacts with the metadata component to perform the management and control functions [6].

The Data warehouse infrastructure

DW infrastructure includes all the foundational elements that enable the architecture to be implemented. The elements of the data warehouse infrastructure may be classified into two categories: operational infrastructure and physical infrastructure. This distinction is important because elements in each category are different in their nature and features compared to those in the other category. The infrastructure includes several elements such as server hardware, operating system, network software, database software, the LAN and WAN, vendor tools for every architectural component, people, procedures, and training [6].

Discussion

In disease management programs, healthcare organization requires data analysis to demonstrate the effectiveness of the treatment plan, report the results of medical interventions and provide the best and highest quality of care along with reducing cost that is resolved with developing and implementing appropriate and successful data warehouse. Since data warehouse with integrated data and powerful analytical capabilities, enables the users to identify patients at risk, Monitor the progression of disease, evaluate the programs and complete the financial analysis. The function of DW is to storage data in historical and multidimensional format which causes the history of patient's episodes of care is available and clinicians using it can analyze the past condition of patient and relate the driven information to current condition of patient and predict the future of her/his disease. Even as needed they can follow disease progress by daily, weekly or monthly and analyze the disease process. In addition the clinicians can also compare the various episodes of disease to improve their performance effectiveness and gather the result of medical intervention to compare them with standards in order to assess their performance.

Among the centers that have used this system can be referred to Bals Infectious Diseases Institute in Bucharest, Romania, where used DW to manage the complicate conditions in AIDS management [23]. In

Netherlands, the department of medical informatics of Rotterdam university medical center developed the DW with the name of «AsthmaCritic» for the management of patients with asthma or COPD which can enable the clinicians to select and analyze the electronic data of patients and design the appropriate plan for them. The system also can assess the care plan of clinicians and provide the feedback system to benchmark of their practice against most common practice guidelines and standards [24].

As regards, all of data related to diagnostic and treatment procedures are available in DW; the clinicians can refrain from re-ordering them and save the costs. As well as, the managers can identify the utilization rate of facilities and resources and the factors affecting their utilization rate to planning and allocating the budget. As Athena health care organizations in the United States where developed a DW to achieve betterment and cost-effectiveness in diabetes management as well as, the Blue Cross and Blue Shield companies in U.S.A, Tennessee developed DW to manage heart disease and diabetes in the primary care level [19,25].

Conclusion

Finally, with data warehousing can optimize disease management by decreasing outbreaks, improving patient care and reducing healthcare cost which can enhance quality of patient's life.

References

1- Perry, T. Technology-Driven Outcomes. Health Management Technology 2004; 25(1):40

2-RamickD C Data Warehousing in Disease Management Programs. Journal of Healthcare Information Management 2001; 15 (2).

3- karami M. the better management of disease with data warehousing. The first international conference on telemedicine and electronic health. Shahid beheshti university in medical science, Iran, Tehran1-2 novamber 2006. [article in Persian]

4-Shams K, Farishta M. Data warehousing: toward knowledge management. Topics in Health Information Management 2001; 21(31): 25-2.

5- Silver M, Sakata T, Su H, Dolins SB, Oshea M J. Case study: how to apply data mining techniques in healthcare data warehouse. Journal of Healthcare Information Management 2001; 15.

6- Ponniah p. Data warehousing fundamentals. New York: john Wiley & sons; 2001.

7- Thomas G. The Decision-Driven Warehouse - Attention Warehouse Shoppers: The Output is in the Decision. URL: http://articles.techrepublic.com, 2006.

8- karami M. Managers & Decision Support Systems, homaye Salamat, inpress. [article in Persian]

9-Watson H, Ariyachandra T, Matyska R.J.Jr. Data warehousing stages of growth. Information Systems Management 2001; 18(3): 42-50.

10- Padilla Jr R. Beginning and planning your data-warehousing project. Available from: URL: http://articles.techrepublic.com.com/5100-6329_11-5155894.html, 2004.

11-Verma R, Harpen Y. Life cycle of data warehousing project in healthcare. Journal of Healthcare Information Management 2001; 15.

12- Mattison R. Data warehousing: strategies technologies and techniques. MC Grow-Hill: Wiley; 1996.

13- Karami M, Setayesh M. The role of data warehouse to Empower health care organizations congress of Hospital and Information Technology, Iran, Kermanshah, 12 January 2005. [article in Persian]

14- Ladaga, J. Let business goals drive your data warehouse effort. Health Management Technology 1995; 16(11): 26.

15- Kaelin M. Resources for designing, planning, and implementing a data warehouse strategy. 2004; Available from: URL: http://articles.techrepublic.com.com/5100-6228_11-5479908.html, [cited 2007 Nov 12]

16- Vengugopal V. Fang J. Bhat S. Performance Driven Data Integration: An Optimized Approach. 2007; Available from: URL: http://articles.techrepublic.com. [cited 2007 Nov 12]

17-Fuller A. Optimizing searches in an OLTP database. Available from: URL: http://articles.techrepublic.com, 2005.

18- karami M, Piri Z. from clinical coding to knowledge management in health care organization. Health Information Management 2007; 4 (2): 239-248. [article in Persian]

.19- Kevin C W. Data warehouse techniques for outcomes management. Health Management Technology 1999; 20 (10) : 16-17.

20- Dilorio F, Abolafia J. The Electronic project: effectively using Meta data throughout the project life cycle. 2007; URL: http://articles.techrepublic.com. [cited 2007 Nov 12]

21- karami M. Application of data mining in clinical decision making. The first international conference on telemedicine and electronic health. Shahid Beheshti University in medical science, Iran, Tehran,1-2 novamber 2006. [article in Persian]

22- karami M. Application of Data Mining and Text Mining in agility of healthcare organization Health Management 2007; 10 (30): 15-30. [article in Persian]

23- Bejenariu M. Iancu M. Iliuta O. Nistor D. Clinical Data Mining Decision Support Tool in AIDS Management. 2003; available at: http://www.mrc.ac.za/conference/helina. [cited 2007 Nov 12]

24-Krohn R. Data Analytics throughout the Healthcare Enterprise Journal of Healthcare Information Management 2001; 18 (2):17

25–Prather J.C, et al. Medical data mining: knowledge discovery in a clinical data warehouse. 2005;Available from: URL: www.dci.mc.duke.edu/PDF_Files/Data%20Mining.pdf. [cited2007 Nov 19]