



WestminsterResearch

<http://www.wmin.ac.uk/westminsterresearch>

An OLAP-enabled software environment for modelling patient flow.

Christos Vasilakis
Elia El-Darzi
Panagiotis Chountas

Harrow School of Computer Science

Copyright © [2006] IEEE. Reprinted from the Proceedings of the 3rd International IEEE Conference on Intelligent Systems, pp. 261-266.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Westminster's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners. Users are permitted to download and/or print one copy for non-commercial private study or research. Further distribution and any use of material from within this archive for profit-making enterprises or for commercial gain is strictly forbidden.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of the University of Westminster Eprints (<http://www.wmin.ac.uk/westminsterresearch>).

In case of abuse or copyright appearing without permission e-mail watts@wmin.ac.uk.

An OLAP-Enabled Software Environment for Modeling Patient Flow

Christos Vasilakis, Elia El-Darzi, and Panagiotis Chountas

Abstract—On-Line Analytical Processing (OLAP) tools use multidimensional views to provide quick access to information. They have become the de facto standard in the business world for analytical databases. In health care, care givers and managers could benefit from being able to perform interactive data exploration, ad-hoc analysis and possibly discover hidden trends and patterns in health data. However, health data have unique characteristics that distinguish them from common business examples, an aspect that makes the direct adaptation of the already established business oriented solutions difficult. In this paper we report the development of an OLAP system for analyzing hospital discharge data and for modeling hospital length of stay.

Index Terms— On-line analytical processing (OLAP), health systems, patient flow.

I. INTRODUCTION

It has become a cliché that there are no good or bad decisions, rather decisions based on good or bad information. At the dawn of the 21st century and in the era of information superhighways everyone is well aware of the importance of the right information, in the right form, at the right time, and to the right people. In the last decades we have witnessed tremendous advances in every technological field that is related to the creation, processing and dissemination of information. These advances however, were mainly driven by the corporate world and consequently organizations of the public sector were slow in adopting the new technologies. This phenomenon is also evident in the health care sector of the developed countries.

The general consensus is that there is a strong need for rigorous examination of all the facts concerning health care issues at local, national or global level. Demographical problems and more specifically the ageing population, coupled with tighter budgets have increased the pressure on finding ways to provide more efficient and effective services for ever increasing numbers of patients. As a result, everybody involved in the provision of health care, from care givers to managers and from social-service workers to local authorities, is increasingly required to operate within an environment of performance indicators and strict financial regulations. In most cases however, the cutting edge technologies available to the decision makers of the business sector are not readily available to their counterparts in the health care sector.

This paper focuses on the analysis of hospital length of stay (LoS) and bed occupancy in geriatric hospital departments. A recent beds enquiry in the UK showed that two thirds of hospital beds are occupied by patients aged 65

and over [1]. This phenomenon is not only attributed to the higher admission rate (289 per thousand population for the 65+ age group as opposed to 94 per thousand population for the 15 to 64 age group) but also to the almost twice as long average LoS of this group of patients. Hence, we believe the provision of tools to aid in the analysis of hospital LoS and bed occupancy is critical to the management of these patients and to the allocation of health resources.

The rest of the paper is organized as follows. In section 2 we describe the relevant information technology advances and review the related literature. In Section 3 we briefly describe relevant methods for measuring and modeling hospital LoS of patients. In Section 4 we present the relational and multidimensional data models for LoS and bed occupancy analysis and give examples of OLAP queries. In Section 5 we summarize the contribution of this paper and identify areas for future research and development.

II. DECISION SUPPORT SYSTEMS, DATA WAREHOUSING AND OLAP

Decision support systems is a category of information systems that supports decision making regardless of the type of internal model that is used [2]. Smith [3] argues that the deployment of decision support systems in health organizations is necessary since they can satisfy the need for accurate summaries and aggregations which is easy to use. For a detailed account of decision support systems in health services management see Forte and Cropper [4].

Decision support systems pose some different requirements on database technology than the traditional on-line transaction processing (OLTP) applications [5] that mainly automate day-to-day, clerical data processing tasks. Examples of health care OLTP applications include the electronic patient record systems (EPR) and the patient activity tracking systems (admissions, discharges, transfers).

Decision support systems require historical, summarized and consolidated data from many different sources scattered throughout the organization and quite often, outside the limits of it. The data warehousing approach is an attempt to overcome these problems by integrating data from operational database and from external sources into one single data repository called the data warehouse. A data warehouse is a subject oriented, integrated, non-volatile and time variant collection of data in support of management's decisions [6]. The fundamental function of data integration makes the adoption of the data warehousing philosophy extremely important for supporting decision making in the field of health care. A reason for this is the existence of different medical and coding standards, many of which are incompatible or

All authors are with the Harrow School of Computer Science, University of Westminster, Northwick Park, Harrow Road, Harrow HA1 3TP, UK. (telephone: +44-20-7911-5000, e-mail: {C.M.Vasilakis, eldarze, countp}@westminster.ac.uk).

require very careful interpretation [7]. The adoption of data warehousing techniques is further necessitated by the fact that the results must usually be delivered to a wide range of stakeholders: care givers, managers, regulators, social-service workers, health economists, researchers, patients and carers etc.

The multidimensional data model that emerged in the past decade has been successful in analyzing large volumes of data for decision making purposes [8]. Multidimensional databases logically view data as multidimensional data cubes that are suited for data analysis. On-Line Analytical Processing or OLAP [9], is a software technology that takes advantage of this data model by enabling analysts, managers and executives to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information that has been transformed from raw data to reflect the real dimensionality of the organization as understood by the user [10]. Data are seen either as facts with associated numerical measures or as textual dimensions that characterize these facts. Dimensions are usually organized as hierarchies. Typical OLAP operations include the aggregation and de-aggregation of data along a dimension (roll-up and drill-down), the selection of specific parts of the cube (slicing), the reorientation of the dimensions in the multidimensional view of data on the screen (pivoting) and the displaying of values from one dimension within another one (nesting) [11]. OLAP is essentially a tool for browsing data stored in the data warehouse and it does not necessarily imply that the underlying physical model of the database is multidimensional.

A number of health data warehousing and OLAP efforts have been reported in the literature. Berndt *et al* [12] reports the building and operation of a comprehensive data warehouse that contains health care data from the State of Florida. It supports health assessments of communities throughout the state and other areas of public health decision making. The suitability and value of hospital discharge transactions as a data warehouse component is also highlighted [13]. Ewen *et al* [14] and Scheese [15] critically discuss the business case for the need of a clinical data warehouse. Hristovski *et al* [16] reports the development of an outpatient data warehouse at a national level and the employment of OLAP for building decision support systems in the domain of public health care. Ebidia *et al* [17] discusses the process of creating OLAP capacity from data generated by the electronic patient record (EPR), as do Ledbetter and Morgan [18], which suggests that data acquired by an EPR system should be remodeled in a data warehouse before being employed by a clinical support system. Zilli [19] discusses the role of OLAP in an outpatient health care statistics data warehouse. Gordon *et al* [20] provides examples of potential OLAP applications for care improvement projects from the perspective of the emergency department. On a different note, Pedersen and Jensen [21, 22] notably expose the limitations of the current data models and technologies in modeling complex clinical data and propose alternatives.

Furthermore, data warehousing and OLAP have recently been used in conjunction with data mining techniques. Isken and Rajagopalan [23] describes such an effort for

classifying patients into classes with similar consumption patterns of hospital resources. Silver *et al* [24] demonstrates how data mining algorithms can be used to identify regions in the data with irregular behavior and then how OLAP can complement data mining by allowing the users to compare the different regions in greater detail.

III. MEASURING AND MODELING THE FLOW OF PATIENTS

The inefficiencies of traditional methods in describing patient activity with skewed LoS distributions is well documented in the literature [25]. Simple LoS averages can offer indications but cannot accurately describe the process of care in such hospital departments as geriatric or psychiatric [26]. The complicating factor is the presence of patients with considerably longer LoS than others, in many cases in the order of months. Consequently decisions on resource allocation and patient management that are based on such measures are often suboptimal [27].

Alternative methodologies have been developed to overcome this problem [28]. For instance, McClean and Millard [29] have modeled the LoS of geriatric patients by using a two-term mixed exponential distribution. More generally, the observation that the LoS can be described in terms of mixed exponential equations have lead to the development of a flow model for modeling patient activity [30]. The different streams of patient flow can be classified as short stay (usually measured in days), medium stay (measured in weeks), and long stay (measured in months or years). The main advantage of the existing software application that implements this model is that only a bed census is required for estimating the input parameters [31]. Discrete event simulation models that have been developed based on the multi-stage nature of patient flow, have further extended the capabilities of the mathematical models by incorporating the stochastic nature of the system under study and the cascading effect of bed blockage in measuring the performance of alternative policies [32].

Hospital LoS however, is not the only parameter to consideration in describing patient activity in hospital departments [33]. Patterns of admissions and discharges and of overall occupancy play a major part in understanding the system. Weekly and seasonal variations are well reported and account for major disruptions and bed crises [34]. For example, a major seasonal variation in the pattern of admissions and discharges is the cause behind the well publicized winter bed crisis, a cyclical phenomenon that used to appear in British hospitals, two or three weeks after Christmas. The rise in admissions has often been proposed as a possible explanation, however, analysis of data from a teaching hospital with bed shortages suggested delays in discharging elderly patients as a possible alternative explanation [35].

IV. THE APPLICATION OF OLAP

There are two different types of analysis that the proposed system facilitates, hospital LoS and bed occupancy. From a data modeling point of view they are different because the granularity of the fact table differs. For hospital LoS analysis, details of each spell (defined here as the period commencing with admission of a patient

to hospital and ending on discharge) are required for correctly calculating LoS statistics and for feeding the flow and simulation models with the appropriate input data. Thus, a transaction-based fact table is proposed. For bed occupancy analysis, a snapshot fact table identified by the combination of the dimensional keys is proposed as daily occupancy is perceived as an accumulating fact.

A. Data Models for Transaction-Based Analysis

In the star schema of the transaction-based data model, patient spell is considered as the fact and LoS as the measure, Fig. 3. The primary key in the fact table “LOS” is `spell_id` since the granularity of this table has to allow for LoS calculations. The actual date in the fact table has been substituted by two “time codes” (`admission_time_code` and `discharge_time_code`), recommended practice in designing the fact table in a data warehouse [36].

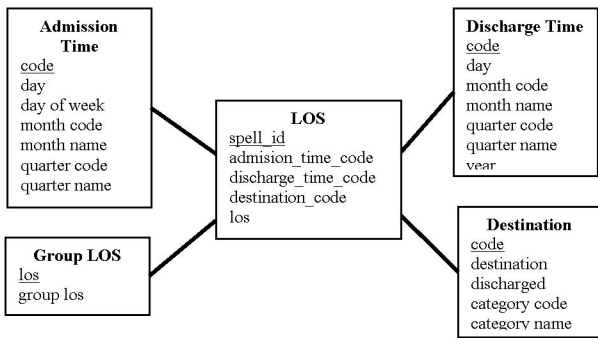


Fig. 3. Star schema for the transaction-based data model

The dimension tables “Discharge Time”, “Admission Time”, “Destination”, and “LOS” have a one-to-many relationship with the fact table. The “Discharge Time” and “Admission Time” tables for the time dimensions have the attributes `month_code`, `month_name`, `quarter_code`, `quarter_name`, and `year`. This schema can be easily converted into a snowflake schema, if needed, by normalizing the three un-normalized tables “Admission Time”, “Discharge Time”, and “Destination”.

The attribute `los` in the “LOS” fact table, apart from being the metric of the star schema, is also the foreign key for linking the fact table with the “Group LOS” dimension table. This dimension can be used for stratifying the daily bed occupancy and for generating frequency distributions of LoS. Typically, each finished spell can be grouped according to LoS into several groups that are based on clinical and managerial judgment.

If the data are stored in a relational database (ROLAP architecture), then the above dimensional model is sufficient. Once this schema has been implemented and populated with data, SQL statements can be used to extract information from it.

In multidimensional databases (MOLAP architecture) data are stored in n -dimensional arrays [11]. For designing purposes, the “data cube” is commonly accepted as the underlying logical construct to conceptualize these multidimensional databases [37]. Adapting the model proposed by Thomas and Datta [37], a data cube can be defined as a 5-tuple $\langle C, A, f, d, O \rangle$ where C is a set of

characteristics, A is a set of attributes, f is a set of one-to-one mappings between a set of attributes to each characteristic, d is a Boolean-valued function that partitions C into dimensions (D) and measures (M), and O is a set of partial orders. The sixth tuple in the Thomas and Datta model [37], L , is omitted here as it is used to instantiate the cube and does not influence its design. Data cube LOS is then defined as follows:

$C = \{\text{admission_time, discharge_time, destination, group_los, spell_los}\};$

$d(\text{admission_time})=1, d(\text{discharge_time})=1,$
 $d(\text{destination})=1, d(\text{group_los})=1, \text{ and } d(\text{los})=0;$

$D = \{\text{adm_day, adm_week, adm_month, adm_quarter, adm_year, dis_day, dis_week, dis_month, dis_quarter, dis_year, destination, discharged, destination_category, group_los}\},$

$M = \{\text{spell_id, los}\}$ and thus $A = D \cup M;$

$f(\text{admission_time}) = \{\text{adm_day, adm_week, adm_month, adm_quarter, adm_year}\},$

$f(\text{discharge_time}) = \{\text{dis_day, dis_week, dis_month, dis_quarter, dis_year}\},$

$f(\text{destination}) = \{\text{destination, discharged, destination_category}\},$

$f(\text{group_los}) = \{\text{group_los, los}\}$ and

$f(\text{spell_los}) = \{\text{spell_id, los}\};$

$O_{\text{admission_time}} = \{\langle \text{adm_day, adm_month} \rangle, \langle \text{adm_month, adm_quarter} \rangle, \langle \text{adm_quarter, adm_year} \rangle, \langle \text{adm_week, adm_year} \rangle\},$

$O_{\text{discharge_time}} = \{\langle \text{dis_day, dis_month} \rangle, \langle \text{dis_month, dis_quarter} \rangle, \langle \text{dis_quarter, dis_year} \rangle, \langle \text{dis_week, dis_year} \rangle\},$

$O_{\text{destination}} = \{\langle \text{destination, destination_category} \rangle, \langle \text{destination_category, discharged} \rangle\},$

$O_{\text{group_los}} = \{\langle \text{los, group_los} \rangle\}$ and,

$O_{\text{spell_los}} = \{\}$

A segment of the data cube is graphically illustrated in Fig. 4. For presentation reasons, only three of the dimensions are included, “Discharge Time”, “Location”, and “Destination”. Each cell of the cube represents the “Spell LOS” set of measures corresponding to the value of the three dimensions.

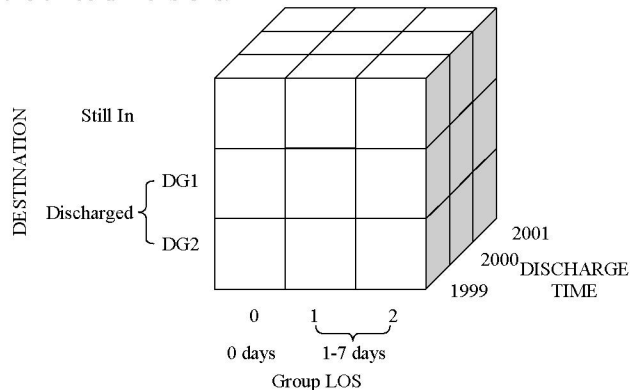


Fig. 4. Instance of the data cube for the transaction-based data model

Several decision support queries can be performed on the LOS cube by using a combination of OLAP operators such as slicing, pivoting, and aggregation. OLAP algebra

operators can be used to formally define these operations.

Suppose the user would like to display the average LoS per discharge destination i.e. the metric `los` needs to be averaged over the destination level of the “Destination” dimension. Thus, assuming that Γ is the aggregation operator [37], and that AVERAGE is the standard SQL operator, this query can be answered by the following operation:

$$\Gamma_{[AVERAGE, \{destination\}, los]}(LOS) = C_{RESULT}$$

The above query can also be considered as a roll-up operation, from the lowest level of the hierarchy to one level up.

Suppose the user wants to roll-up to a higher level in a dimension, for instance, to find the total number of bed days in each quarter of admission date. Such a query requires two aggregations, firstly `adm_day` to `adm_month` and secondly `adm_month` to `adm_quarter`. For the first aggregation, we transform `adm_month` to a measure so that it is retained for the second aggregation (operator Ψ). For the second aggregation, we transform `adm_month` back into a dimension (operator Φ ; for definitions of both operators see [37]). The standard SQL function SUM can be used to calculate the total number of bed days. This aggregation can be expressed as follows:

$$\Gamma_{[SUM, \{adm_quarter\}, los]}(\Phi_{adm_quarter, admission_time} \cdot (\Gamma_{[SUM, \{adm_month\}, los]}(\Psi_{adm_quarter, admission_time} \cdot (LOS)))) = C_{RESULT}$$

A midnight bed census can be generated by restricting the values of the cube for a given date:

$$\sum_{Admission_Time \leq date \wedge Discharge_Time \geq date} \cdot (LOS) = C_{RESULT}$$

C_{RESULT} is a 4-dimensional array that contains LoS observations at the level appropriate for further analysis. Sampling issues related to LoS observations can easily be addressed by, for example, querying only for discharged patients. Alternatively, the following query can be used to create the maximum day bed census on a given date:

$$\sum_{Admission_Time \leq date \wedge Discharge_Time > date} \cdot (LOS) = C_{RESULT}$$

B. Data Model for Snapshot-Based Analysis

Having described the transaction-based data model that facilitates LoS analysis, a model specifically designed for occupancy analysis is detailed in this section. A variation of the inventory periodic snapshot-based design is used. This design, proposed by Kimball and Ross [36], measures the inventory levels of a product and places them as separate rows in a fact table while using the standard metric “Quantity on Hand”. In this case, patients are considered to be the products and occupancy the inventory. Furthermore,

because we are interested in the time-series of daily admissions and discharges, instead of using the standard metric “Occupancy”, we use the metrics “Admissions” and “Discharges”. The daily occupancy is then reconstructed by using suitable operations. The advantage of not storing daily occupancy data for every key combination is that the resulting snapshot table is not as dense as it would have otherwise been.

The star schema contains one fact and three dimension entities (Fig. 5). The fact table has composite primary key comprising the attributes `time_code`, `destination_code` and `group_los_code`. These attributes are also foreign keys, linking the fact and dimension tables with one-to-many relationships. The field `time_code` represents the date of the snapshot.

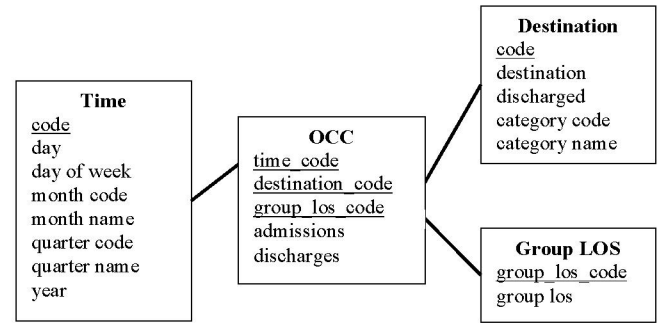


Fig. 5. Star schema for the snapshot-based data model

A date cube based on the above schema, named OCC, is defined as follows:

$C = \{time, destination, group_los, adm_dis\};$
 $d(time) = 1, d(destination) = 1, d(group_los) = 1,$ and $d(adm_dis) = 0;$
 $D = \{day, week, month, quarter, year, destination, discharged, destination_category, group_los\},$
 $M = \{admissions, discharges\};$
 $f(time) = \{day, week, month, quarter, year\},$
 $f(destination) = \{destination, discharged, destination_category\},$
 $f(group_los) = \{group_los\}$ and $f(adm_dis) = \{admissions, discharges\};$
 $O_{time} = \{\langle day, month \rangle, \langle month, quarter \rangle, \langle quarter, year \rangle, \langle week, year \rangle\},$
 $O_{destination} = \{\langle destination, destination_category \rangle, \langle destination_category, discharged \rangle\},$
 $O_{group_los} = \{\}$ and, $O_{adm_dis} = \{\}$

Again, OLAP algebra can be employed to formally define various decision support queries. Suppose the user wants to plot the seven-day moving average of admissions and discharges. Let $MA(n)$ represent a function that takes as input n cube cells and returns an attribute corresponding to their average, then this query can be expressed as:

$$(\Gamma_{[MA(7), \{day\}, admissions]}(OCC)) \cup \cdot (\Gamma_{[MA(7), \{day\}, discharges]}(OCC)) = C_{RESULT}$$

The following query generates the maximum daily occupancy on day x :

$$(\Gamma_{[SUM, \{day\}, admissions]}(\sum_{day=x} (OCC)) - \\ \cdot (\Gamma_{[SUM, \{day\}, discharges]}(\sum_{day=x-1} (OCC))) = C_{RESULT}$$

while the following query includes `group_los` in the resulting cube:

$$(\Gamma_{[SUM, \{day, group_los\}, admissions]}(\sum_{day=x} (OCC)) - \\ \cdot (\Gamma_{[SUM, \{day, group_los\}, discharges]}(\sum_{day=x-1} (OCC))) \\ = C_{RESULT}$$

The implementation of the dimensional model can take place in any commercial relational database management system (RDBMS) such as *Oracle* or *SQL Server 2000*, or an OLAP server that supports the ROLAP architecture. In either case, standard SQL can be used to perform the analytical queries. The data cube can be implemented using an OLAP server that supports the MOLAP architecture such as Microsoft's *Analysis Services*. The queries can then be written in *Multidimensional Expressions* (MDX), a language specifically designed for OLAP analysis [38].

V. DISCUSSION AND FURTHER DEVELOPMENTS

The presence of an OLAP-enabled data warehouse benefits the application of decision support systems within any organization. Less time is needed for collecting, understanding, and cleaning data for analytical purposes. More accurate and sophisticated models can be developed to address a greater number of problems and functions.

In this paper we described relational and multidimensional data models that can be used in an OLAP-enabled data warehouse environment. A transaction-based model was proposed for facilitating LoS analysis and a snapshot-based for admissions, discharges and bed occupancy. We used the model proposed by Thomas and Datta [37] to describe the multidimensional data model and the OLAP algebra to formally describe OLAP queries.

Data analysis and decision making in health care can benefit enormously by the quickness and easiness of data exploration and by the active participation of the various stakeholders in the actual analysis. Hospital administrators can perform time series analysis of admissions, discharges and bed occupancy either overall or grouped by specialties and/or discharge destination, gender, source of admission etc. Regional planners can compare the performance of the various hospital units in the region and drill down to individual units when variations are observed. Researchers and modelers can benefit by speeding up the lifecycle of studies requiring the development of analytical models. These features, coupled with the capability of integrating data from disparate sources (hospital and nationwide databases, world wide web sources) and the application of data cleansing and data quality procedures, can enhance problem understanding, model building and eventually decision making in health care.

From a model development point of view, analyzing

model output with OLAP tools could also be very beneficial. Typically, analytical models such as simulation can generate massive volumes of data that often require significant processing before mean estimations of the output parameters can be calculated. What-if scenarios applied to a model can be considered as an additional dimension and together with the values of the output parameters can be regarded as the facts of a "Model Output" data warehouse. Such a development could benefit the data-intensive simulation models that were briefly described in this paper. The ultimate objective is to fully integrate an OLAP-enabled data warehouse with the analytical models, both for input and output data analysis.

VI. ACKNOWLEDGMENT

The authors gratefully acknowledge the advisement of Professor Peter H Millard and the useful comments of the anonymous referees.

REFERENCES

- [1] "Shaping the Future NHS: Long Term Planning for Hospitals and Related Services", Department of Health, London, 2000.
- [2] V. Dhar and R. M. Stein, *Intelligent Decision Support Methods: The science of knowledge work*. New Jersey: Prentice Hall, 1996.
- [3] J. Smith, *Health Management Information Systems: A handbook for decision makers*. Buckingham: Open University Press, 2000.
- [4] P. Forte and S. Cropper, "Enhancing health services management," *Health service management*, P. Forte and S. Cropper, Eds. Buckingham: Open University Press, 1997.
- [5] S. Chaudhuri and U. Dayal, "Data Warehousing and OLAP for Decision Support," presented at ACM SIGMOD Int. Conf on Management of Data, Tuscon, Arizona, 1997.
- [6] W. H. Inmon, *Building the Data Warehouse*. New York: John Wiley & Sons, 1996.
- [7] D. J. Berndt, J. W. Fisher, A. R. Hevner, and J. Studnicki, "Healthcare Data Warehousing and Quality Assurance," *Computer*, vol. 34, pp. 56-65, 2001.
- [8] T. B. Pedersen and C. S. Jensen, "Multidimensional Database Technology," *Computer*, vol. 34, pp. 40-46, 2001.
- [9] E. F. Codd, S. B. Codd, and C. T. Salley, *Providing OLAP to User-Analysts: An IT Mandate*. Sunnyvale, California: E.F. Codd & Associates, 1993.
- [10] O. Council, "OLAP and OLAP Server Definitions," in *OLAP Council*: OLAP Council, 2002.
- [11] P. Vassiliadis and T. Sellis, "A survey of logical models for OLAP databases," *ACM SIGMOD Record*, vol. 28, pp. 64-69, 1999.
- [12] D. J. Berndt, A. R. Hevner, and J. Studnicki, "CATCH/IT: A Data Warehouse to Support Comprehensive Assessment for Tracking Community Health," presented at Annual Fall Symposium, 1998.

- [13] D. J. Berndt, A. R. Hevner, and J. Studnicki, "Hospital Discharge Transactions: A Data Warehouse Component," presented at Proceedings of the Hawaii International Conference on System Sciences, Hawaii, 2000.
- [14] E. F. Ewen, C. E. Medsker, and L. E. Dusterhoft, "Data Warehousing in an Integrated Health System; Building the Business Case," presented at DOLAP '98, Washington DC, 1998.
- [15] R. Scheese, "Data warehousing as a healthcare business solution," *Healthcare Financial Management*, vol. 52, pp. 56-59, 1998.
- [16] D. Hristovski, M. Rogac, and M. Markota, "Using data warehousing and OLAP in public health care," presented at Annual Symposium, 2000.
- [17] A. Ebidia, C. Mulder, B. Tripp, and M. W. Morgan, "Getting data out of the electronic patient record: critical steps in building a data warehouse for decision support," presented at Annual Symposium, 1999.
- [18] C. S. Ledbetter and M. W. Morgan, "Toward best practice: leveraging the electronic patient record as a clinical data warehouse," *Journal of Healthcare Information Management*, vol. 15, pp. 119-131, 2001.
- [19] D. Zilli, "Outpatient health care statistics data warehouse--implementation," *Studies in Health Technology and Informatics*, vol. 68, pp. 50-54, 1999.
- [20] B. D. Gordon and B. R. Asplin, "Using online analytical processing to manage emergency department operations," *Acad Emerg Med*, vol. 11, pp. 1206-12, 2004.
- [21] T. B. Pedersen and C. S. Jensen, "Research issues in clinical data warehousing," presented at 10th International Conference on Scientific and Statistical Database Management, Capri, Italy, 1998.
- [22] T. B. Pedersen and C. S. Jensen, "Multidimensional data modeling for complex data," presented at 15th International Conference on Data Engineering, Sydney, Australia, 1999.
- [23] M. Isken and B. Rajagopalan, "Data mining to support simulation modeling of patient flow in hospitals," *Journal of Medical Systems*, vol. 26, pp. 179-197, 2002.
- [24] M. Silver, T. Sakata, H. C. Su, C. Herman, S. B. Dolins, and M. J. O'Shea, "Case study: how to apply data mining techniques in a healthcare data warehouse," *Journal of Healthcare Information Management*, vol. 15, pp. 155-164, 2001.
- [25] C. Vasilakis and A. H. Marshall, "Modelling nationwide hospital length of stay: opening the black box," *Journal of the Operational Research Society*, vol. 56, pp. 862-869, 2005.
- [26] P. H. Millard, "Current measures and their defects," in *Modelling hospital resource use: a different approach to the planning and control of health care systems*, P. H. Millard and S. I. McClean, Eds. London: Royal Society of Medicine, 1994, pp. 29-37.
- [27] P. C. Smith and M. Goddard, "Performance management and Operational Research: a marriage made in heaven?," *Journal of Operational Research Society*, vol. 53, pp. 247-255, 2002.
- [28] A. H. Marshall, C. Vasilakis, and E. El-Darzi, "Length of Stay-Based Patient Flow Models: Recent Developments and Future Directions," *Health Care Management Science*, vol. 8, pp. 213-220, 2005.
- [29] S. I. McClean and P. H. Millard, "Patterns of length of stay after admission in geriatric medicine: an event history approach," *The Statistician*, vol. 42, pp. 263-274, 1993.
- [30] G. W. Harrison and P. H. Millard, "Balancing acute and long term care: the mathematics of throughput in departments of geriatric medicine," *Methods of Information in Medicine*, vol. 30, pp. 221-228, 1991.
- [31] S. Wyatt, "The occupancy management and planning system (BOMPS)," *The Lancet*, vol. 345, pp. 243-244, 1995.
- [32] E. El-Darzi, C. Vasilakis, T. Chaussalet, and P. H. Millard, "A Simulation Modelling Approach to Evaluating Length of Stay, Occupancy, Emptiness and Bed Blocking in a Hospital Geriatric Department," *Health Care Management Science*, vol. 1, pp. 143-149, 1998.
- [33] A. H. Marshall and S. I. McClean, "Conditional phase-type distributions for modelling patient length of stay in hospital," *International Transactions in Operational Research*, vol. 10, pp. 565-76, 2003.
- [34] "Winter Report 2000 - 2001", Department of Health, London, 2001.
- [35] C. Vasilakis and E. El-Darzi, "A simulation study of the winter bed crisis," *Health Care Management Science*, vol. 4, pp. 31-36, 2001.
- [36] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*, 2nd ed. New York: John Wiley, 2002.
- [37] H. Thomas and A. Datta, "A conceptual model and algebra for on-line analytical processing in decision support databases," *Information Systems Research*, vol. 12, pp. 83-102, 2001.
- [38] G. Spofford, *MDX Solutions with Microsoft SQL Server Analysis Services*. USA: John Wiley & Sons, 2001.