International Journal of Bio-Science and Bio-Technology
Vol. 2, No. 1, March, 2010

Data Warehouse Approach to Build a Decision-Support Platform for Orthopedics Based on Clinical and Academic Requirements

[1]Sheng-Hui Lin, [2]Yuan-Chii G. Lee, and [3]Chien-Yeh Hsu

[1]Orthopedic department, CHIMEI medical center, No. 901, Zhong Hua Rd, Yongkang City, Tainan County, TAIWAN
[2]Graduate Institute of Medical Informatics, Taipei Medical University, No. 250 Wu-Xing Street, Taipei City, TAIWAN
dantes@mail.chimei.org.tw, ycgl@tmu.edu.tw, cyhsu@tmu.edu.tw

Abstract

The continuous quality improvement has become a trend in the contemporary medical society, and that can be achieved by the specialty database implement. Decision-support system in the academic and clinical aspects are included in the process such continuous quality improvement. The database has its limitation in the decision-support due to deficiency of on-line analytic function. The data warehouse offers the sophisticated function for decision-support processes. However, the implement of data warehouse may face a lot of obstacles, included expensive cost and large personnel. We had previously established a database of orthopedics, which collected the patients' data since 2002. The new system was constructed based on this specialty database, the knowledge architectures was build up via specialists committee and accreditation indicators. The major function was to generate sufficient information for decision-support process in the academic and clinical aspects. The execution efficiency of this system is more effective than database. The unique knowledge architecture can form a distinguishing feature of the department. The cost that saved from personnel and time reduced from reports generation for accreditation is remarkable. The stratification of web-based interface application can be assessed through questionnaires; the outcome is satisfactory as what we previously expected. The sophisticate function of the data warehouse is hard to express in a solitary department of the hospital, especially when they had already owned traditional database. The experience of this system construction can be useful as one option for upgrade of specialty database and a step forward to the goal of the continuous quality improvement.

1. Introduction

1.1. The importance of the specialty database

As the purpose of computer systems for patient data management, the comprehensive information system became the essential component in clinical practices. A lot of specific systems had been developed and published, such as COSTAR by McManus[1] and OASYS by Stoodley[2]. The former COSTAR is an integrated, modular system that can be implemented incrementally for medical records, accounts receivable, scheduling, and report generation. The latter OASYS is a kind of medical auditing system based on individual diagnoses or specific therapies.

In the clinical orthopedic practice, enormous patient amount and fast admission turnover result in overloading of medical staffs. Decision-making process in such practice needs abundant information from variable data sources. In order to obtain significant information for decision-making process efficiently, organized data are crucial. The organized data have some characteristics include objectivity, accuracy and analyzable.

1.2. Advantage and disadvantage of data warehouse

Early in 1994, Ruffin[3] had raised the importance of medical application of data warehouse. Prather [4] had published a paper in annual meeting of American Medical Informatics Association (AMIA) on 1997, which described new medical knowledge could be generated via data warehouse discovering. The medical institutes and groups implement the data warehouse increasingly in the recent year. They had developed several experiences and techniques to implement their data warehouses. Kerkri[5] had developed their data warehouse focus on the epidemiologic domain and named this system as EPIDWARE. EPIDWARE is an integrated system for providing access to a collection of heterogeneous medical information systems.

Data warehouse can make up the deficiency of databases. There are commercial products of data warehouse in the market. Implement of such systems is seldom found in the solitary department, because of their huge scale and expensive cost. The professional information technician is necessary for maintaining the system integrity. Those are disadvantages of data warehouse applied in the orthopedic department.

1.3. Establishment the unique system

We create a unique system for clinical and academic use of orthopedic departments, which possess functions of auto-extraction of patient data source from the hospital information system (HIS) and on-line analytic process (OLAP) of the data-warehouse. Through this system, we can obtain valuable information to perform the decision-support process and collections of specific patients' information in HIS.

2. Materials and methods

2.1. Clinical materials and physical system setup

This system serves as decision-support system that contains essential variations for decision-making process. The evaluation of efficiency and power of decision-support process in this system was designed by system executing analysis and questionnaire of users. That will be described in the following paragraphs. This system enrolled the orthopedic patients' information of a medical center in Tainan since 2002. The patients' information mainly contains admission, operation and discharge electronic healthy records. At the mean time, the overall amount of database was about 36,000 records and was increasing by days.

The system is installed in a computer, which equipped with Intel® Pentium 4 and 2 gigabyte RAM. The operation system is Microsoft® windows XP professional edition.

This computer is set up as a server, which is located in the intranet of the hospital. The complete set of HIS executing programs are installed in this server. The communication between HIS and the system is via intranet. The end user can access this system through any clients in the hospital.

2.2 System architecture

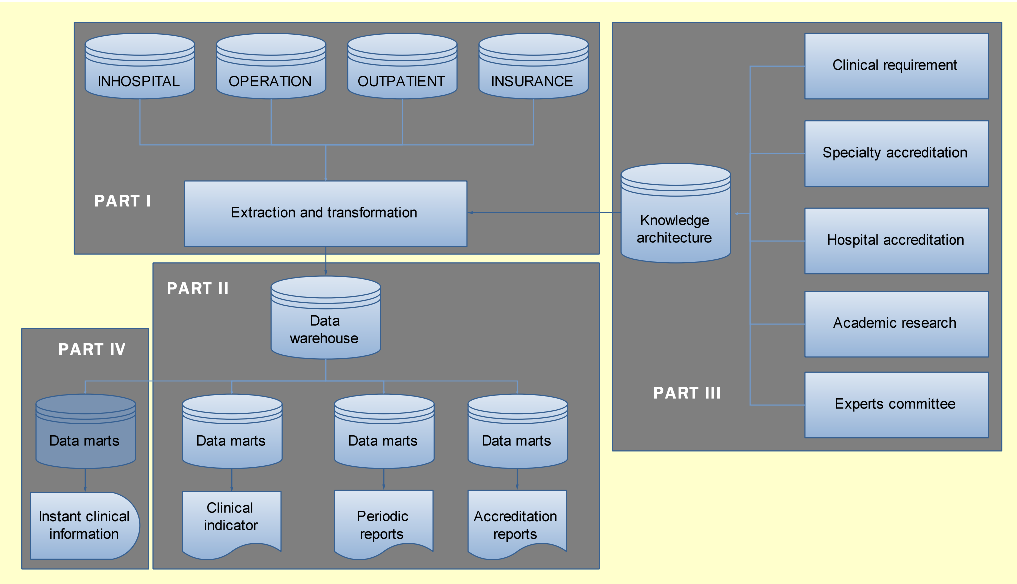

Figure 2-1. Diagram of system architecture

Part I of the system -- The system can extract the daily records from HIS every 5 minutes by a Visual Basic (VB) based subroutine and verifies their consistence to pre-existed data. Through this process, we can ensure the timeliness and accuracy of the database. In the data extraction process we encounter some situations of data error, which included missed entries and conflicted data. Delayed data input and duplicated input by different persons may be the main causes. The medical staff may input the data on the following days after admission, or the operation records may be updated after several days of the admission. Because the data extraction from HIS is occurred on admission date, there was no update can be done in this system while records had been changed beyond the admission date. To correct these invalid data, we design an automatic periodic procedure to extract recent two months data retrospectively every 2 hours. That can renew the out-of-date data and maintained the integrity of database. (Figure 2-2) *Part II of the system* -- The main framework was built by Access®, which possesses four different components – table, form, query and report. (Figure 2-3) The system has three main tables of data that include inpatient, surgery and outpatient source. The tables represent the raw data extracted from HIS, which comprise all fields mapping in databases of HIS in addition to some fields of coding systems. There are several different forms for monitoring and manipulation of data by administrator. *Part III of the system* -- The knowledge architecture was formed in two fashions, one is individual requirements and the other is the conclusions of expert committee. We gathered clinical questions from physician's demand. For example, what is mortality rate in the geriatric hip fracture population of the medical center? The expert committee was hold monthly; it discussed issues and indicators about the research and surgeries of the department. Such information can serve as topics of academic researches and department accreditation data. The questions from individual physician and research topics can be formed as queries of database. The indicative surgeries statistics can be created by specific queries and calculations. Fundamentals of such productions are SQL queries. To form a "WHERE" clause will test the queries results from several times by subspecialty physicians. The knowledge architecture forming process is a continuing task in our department, which was evolved after every expert meeting and new questions, developed in daily practices. *Part IV of the system* –- The final query outcome that includes groups of patients' data and statistics results are demonstrated through web interface. Our physicians can access this web interface via intranet of our hospital. Every physician has their own authority to look up their patient collection and specific patients populations of their researches. The system will generate the statistic reports periodically. Most of web interface are written by ASP and ASP.net. The multiple web pages are connected to a single database located in the web files directory. This database comprises user-requesting data stored in a fragmented fashion.

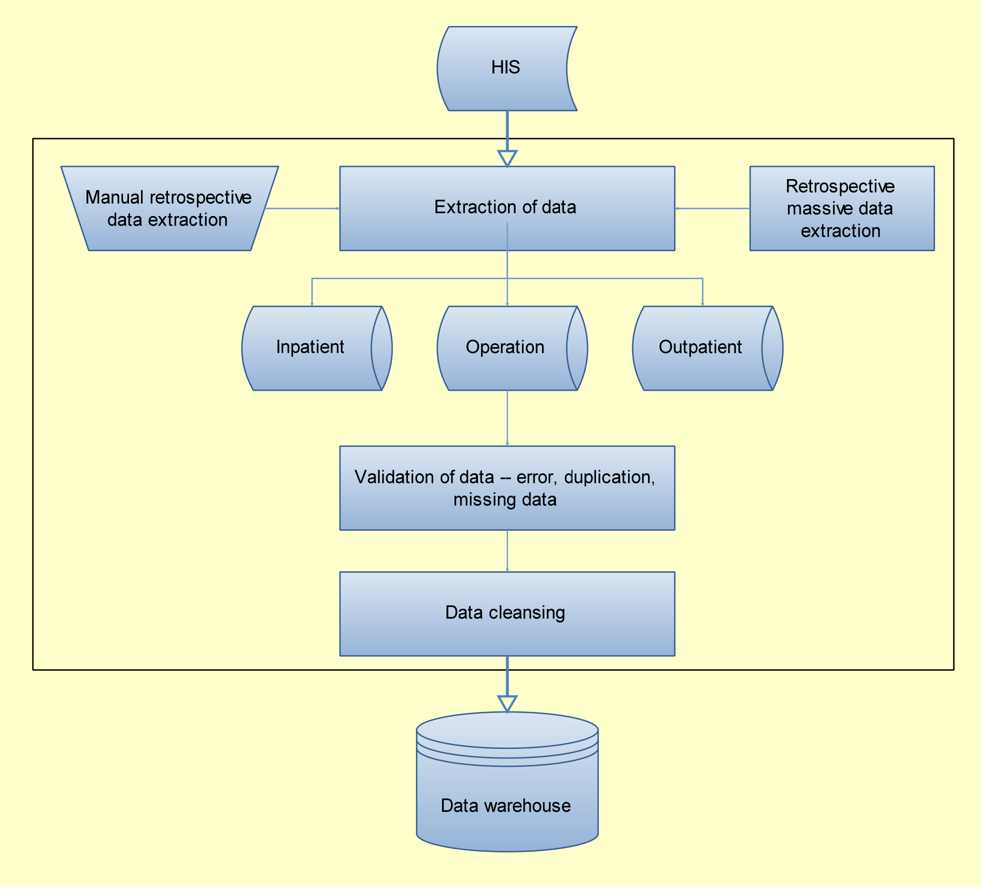

Figure 2-2. Data flow of part I of system

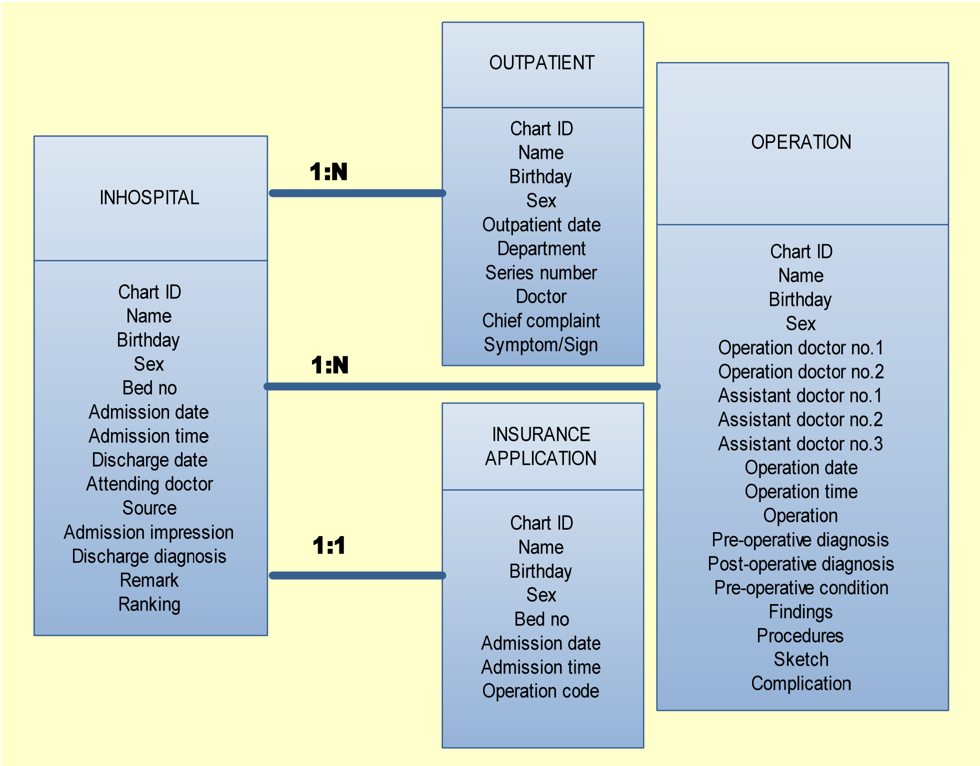

Figure 2-3. Table schema of part II of system

2.3. System evaluation

System evaluation was divided into two aspects – actual executing efficiency and Statistics of invalid data.

Actual executing efficiency is mainly focus on the difference of traditional database query executing and the on-line analytic process of this system. We calculate the executing time of several queries and access error times in the traditional HIS database. The same calculations were applied on this system to reveal their differences.

The invalid data are defined as inconsistent records, missing records and typing error. Such invalid data can not be eliminated by retrospective data extraction. They already exist in the HIS. We adopt specific SQL queries and data mining techniques to generate the periodic report of invalid data and removed them from the system. We compare the statistics before invalid data elimination with that after elimination process to reveal the significant error amount of HIS.

3. Results

3.1. Production

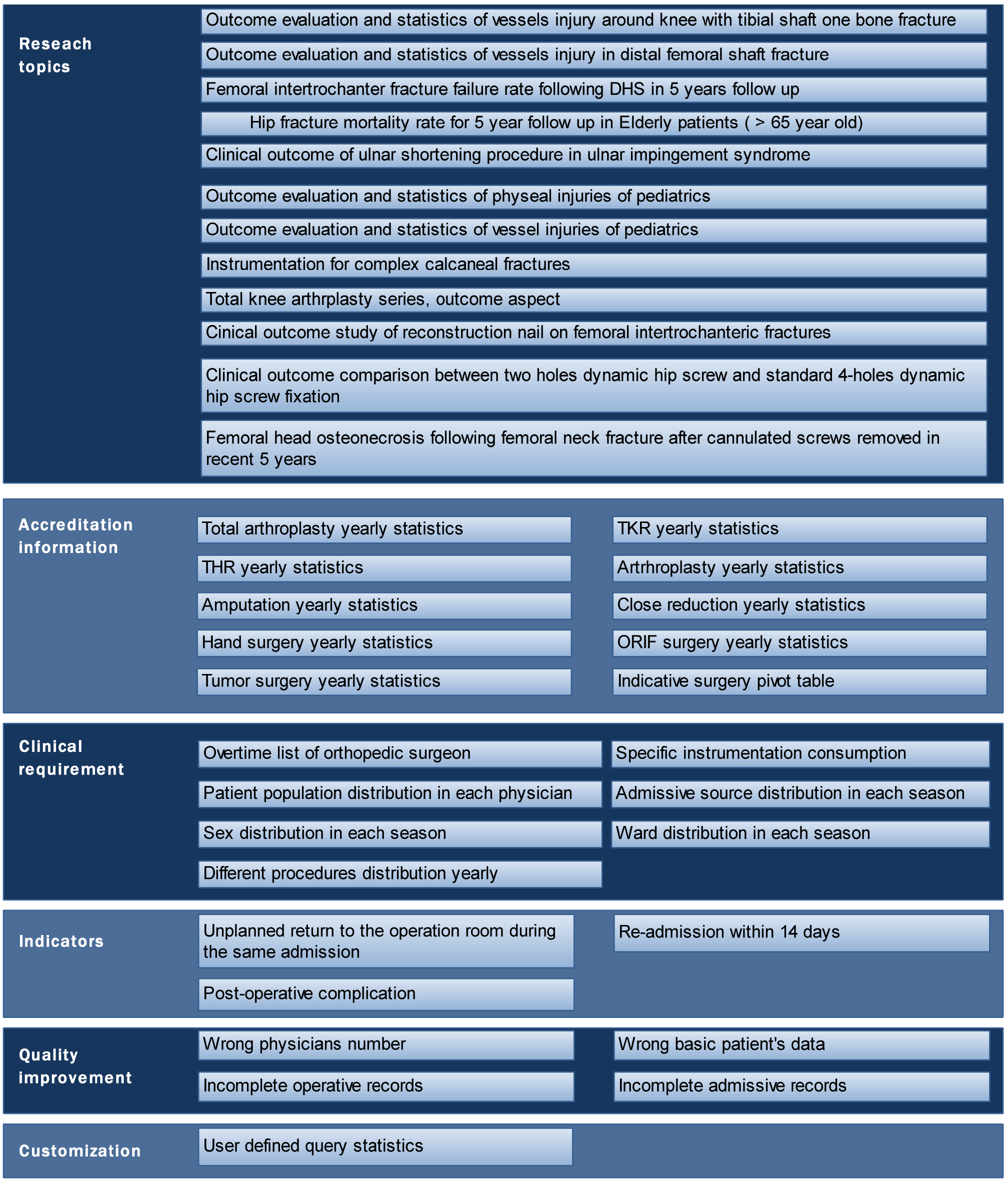

Figure 3-1. Overview of system functions

The production of the system can be presented as reports and forms. (Figure 3-1) The former can be generated by manual or automatic operations. The latter is created as the web-form interface which possessed interaction function and reveal the immediate requested information. The reports of the system can be divided into five aspects: accreditation reports, periodic reports, academic resources and indicators. *Accreditation report* -- Accreditations in the orthopedic department include hospital and specialty accreditation. The core values of hospital accreditation are to provide a safe, effective, patient-centered, timely, efficient and equal healthcare, to establish teamwork to meet community's health need and encouraging hospitals to pursue excellence in their own specialties. The system can generated statistics reports of indicator surgeries through manual or automatic fashion, moreover, the number and type of indicator surgeries also can be appended and changed.

Table 3-1. Statistics of indicative surgeries

INDEX	2002	2003	2004	2005	2006	2007	2008	2009	TOTAL
(1). Total arthroplasty	50	107	342	291	319	417	383	91	**2000**
(2). THR	19	43	147	127	124	129	130	28	**747**
(3). TKR	37	70	216	184	223	310	265	67	**1372**
(4). Arthroplasty	90	232	663	586	620	700	664	177	**3732**
(5). Arthroscopic reconstruction	11	20	48	44	77	65	72	21	**358**
(6). Arthroscopy	44	109	318	325	357	345	260	67	**1825**
Amputation	17	32	104	107	80	61	42	8	**451**
Close reduction	10	27	108	86	79	73	101	16	**500**
DDH	0	6	8	8	6	7	6	0	**41**
HAND surgery	6	8	34	37	41	22	35	6	**189**
ORIF	251	680	2082	2169	2301	2057	1969	454	**11963**
SPINE surgery	33	65	206	163	132	163	142	25	**929**
TUMOR (malignant)	4	9	10	8	16	10	7	6	**70**

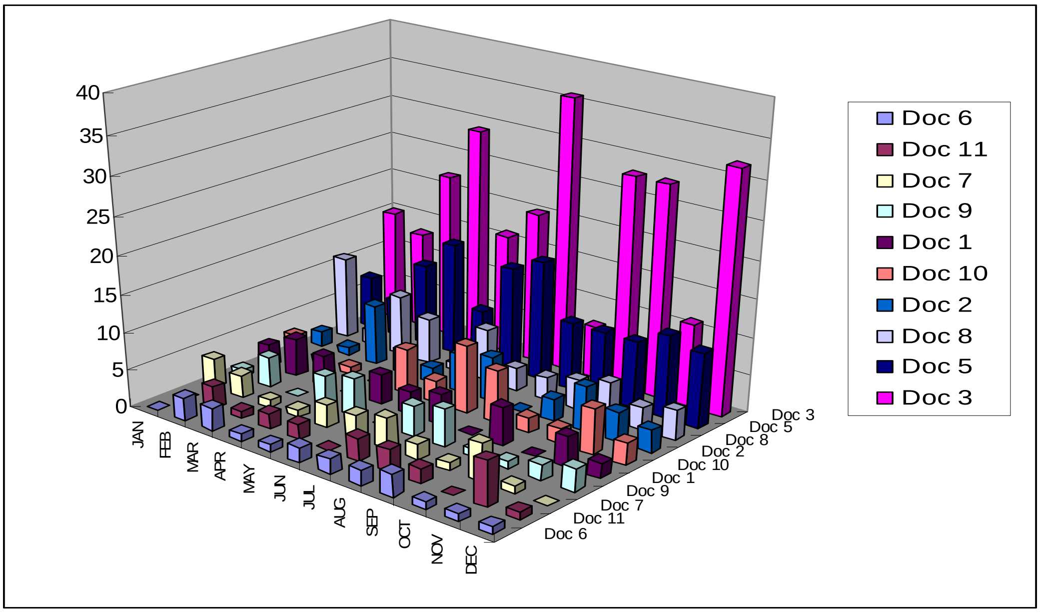

Figure. 3-2. 3D bar chart of patient population of each physician summed monthly in 2005

The indicative surgeries are defined through experts committee, accumulative surgeries amount are presented by pivot table queries, which source is obtained through complex SQL query. Because of single query condition can not encompass the whole indicative surgery; we use multiple filed query and text mining technique to obtain the accurate clinical information. (Table 3-1) *Periodic report* -- The general population of orthopedic patients is enormous and turnover rate is fast. To obtain these information are difficult through HIS, which is not designed for statistics purpose. After mastering this information, the director of department can handle the change of information rapidly. (Figure 3-2) *Academic resource* -- The clinical studies are mainly relied on clinical information of patients. This system can support study proceeding in prospective and retrospective fashion. In the retrospective study, we receive the subjects from researchers and reform their requests to query of Access®. The SQL WHERE clauses are complex and rewritten for several times until output results can meet researchers' requirements. In the prospective study, queries are formatted as the similar form as that formatted in the retrospective study. They can monitor the two groups of patients in the comparison study; do the comparison of pre-defined parameters and accumulation the patient population. (Table 3-2, Figure 3-3)

Table 3-2. Different femoral intertrochanteric fracture implantation

INDEX	2002	2003	2004	2005	2006	2007	2008	2009
2 hole DHS	0	1	7	11	5	6	12	1
Geriatric RECON	0	6	15	15	15	7	15	2
Non 2 hole DHS	14	31	143	132	164	153	144	33

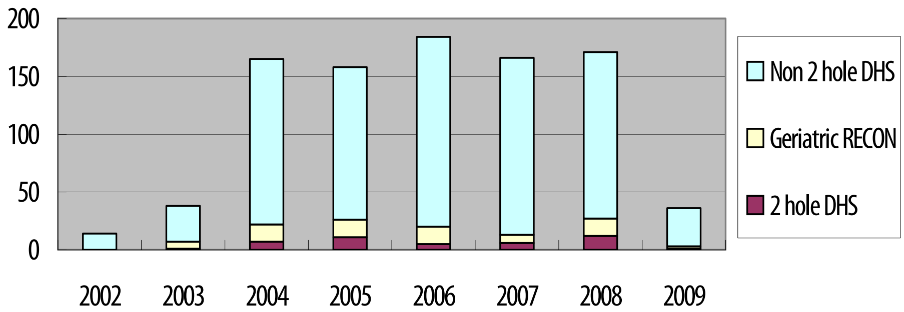

Figure 3-3. Stacked bar chart of different femoral intertrochanteric fracture implantation

Indicator --Since 1999, Taiwan Joint Commission on Hospital Accreditation had introduced the USA's International Quality Indicator Project (IQIP) and developed Taiwan Quality Indicator Project (TQIP). The Commission has been dedicated to promoting quality indicator projects, provided applicable data and information for domestic hospitals. TQIP includes many indicators of different aspects. Manual calculation for indicators is laborious and time-consuming. The easier way to obtain the indicator is extraction of specific data from HIS. The system also offers such functions, which include unplanned/unexpected re-admissions within 14 days, unplanned return to the operating room during the same admission and surgical complication. (Figure 3-4) *Form* -- The web interface can provide functions of immediate access and custom query of clinical information. Application Service Provider (ASP) serves as an easy application tool in the internet use. That faces some problems in the direct connection to the system. Because the system has periodic updating process from HIS, therefore, the end-user's application of ASP will be block by this updating process. That will cause great inconvenient to the end-user. Our solution is to create another database in the directory of web pages. This database only encompasses tables and queries that contained individual physicians' data (not entire department data) and are created according to end-user requests. In the setting period, the system will update this remote database. Through this method, main system processing seldom interrupts web interface usage by end-users.

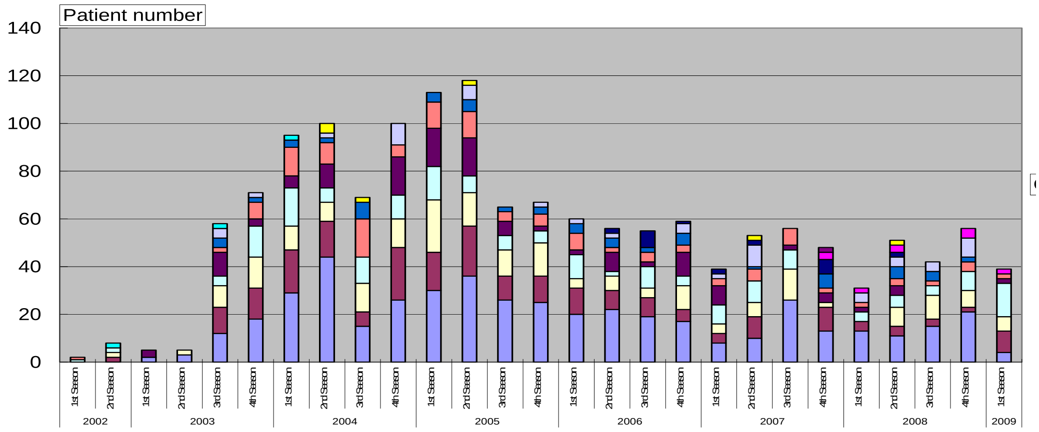

Figure 3-4. Stacked bar chart of unplanned return to the operating room during the same

3.2 Efficiency

Executing efficiency -- In the relational database, time spent in a query will increase as numbers of records increase. (Comparison table each year increment) Generally, time spent in a simple query will not disturb the use of the system possessed orthopedic patients' information of at lease eight years. Pivot table queries (especially multiple query conditions in single column) will spend a lot of time; and affect the whole system execution. Most of results of pivot table queries encompass yearly statistics data. We use the remote database to store results of pivot queries and can be revealed very fast when a query action executes. Results of pivot queries are updated every five hours. That makes a balance between easy of use and accuracy of data. What we compare are mainly in the indicative surgeries. The executing efficiency comparison in time consumption between local machine and remote client is shown in Figure 3-5. *Invalid data detection* -- In the system, we make four queries to detect invalid data, which include wrong physician (not in the orthopedic department), wrong patients' basic information, incomplete operative records and incomplete admissive records. All that invalid data will result in incorrect statistics and reflect some problems in HIS programming and medical behavior. In errors of patients' basic information statistics, the decreasing number reveals improvement of HIS programming or debugging. The incomplete medical records (admissive and operative records) reflect medical staffs can not finish their task in time; may cause some legal problems. (Figure 3-6)

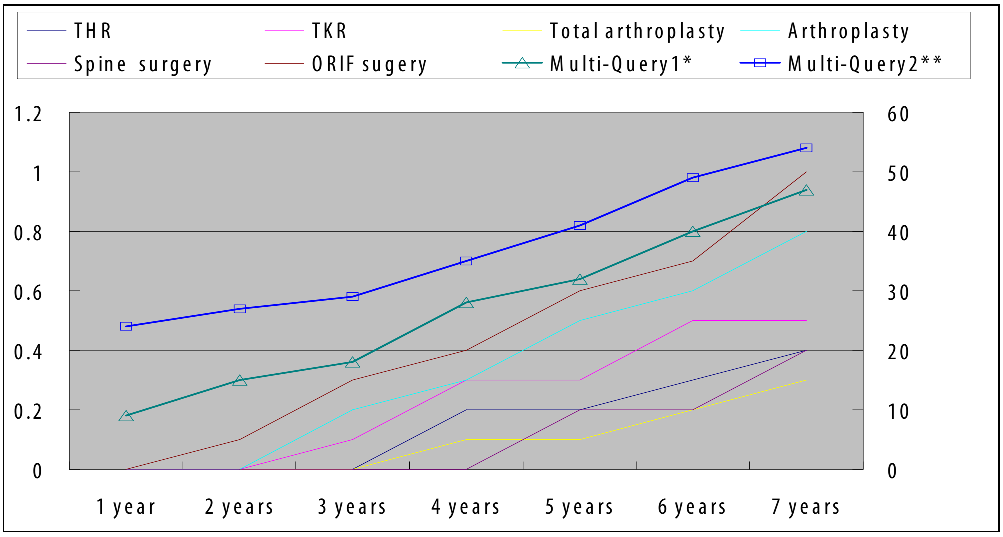

Figure 3-5. Difference(in seconds) between local and remote executing efficiency. Multi-Query1 and Multi-Query 2 show their scale on the right side.

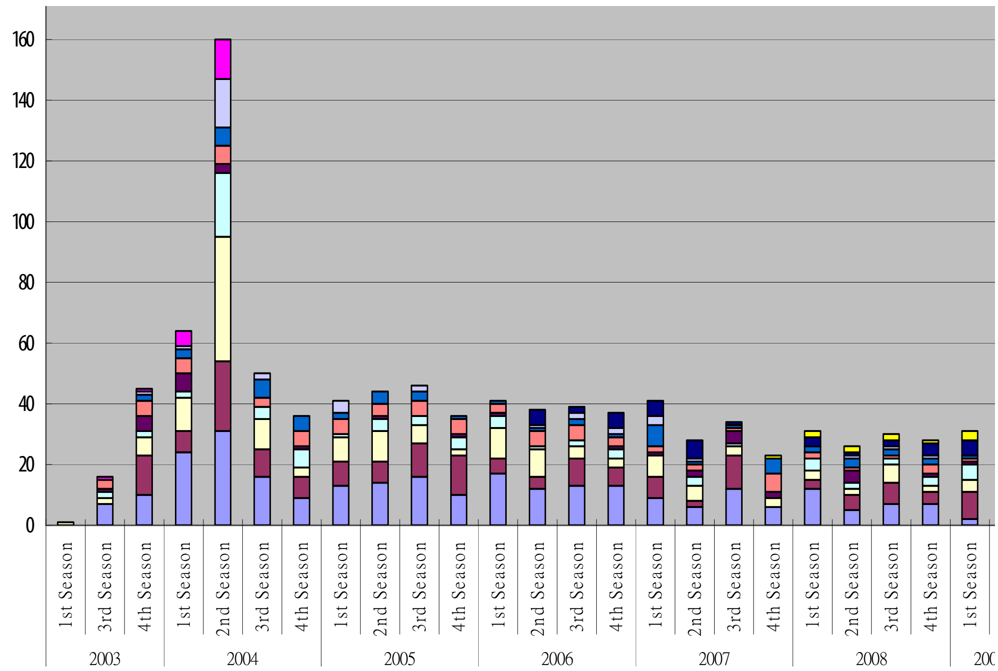

Figure 3-6: Stacked bar chart of number of incomplete operative records. (each physician) The abrupt large number on second season of 2004 is due to massive loss of operative data in HIS.

4. Conclusion and discussions

4.1. Discussions

The system creates various and enormous products to meet users' requests. To achieve such purpose, knowledge architecture plays a crucial role. After the more brain-storming occurs, the more abundant knowledge architecture develops. If we expand the application range to other departments, that will develop a huge knowledge architecture and possess versatile functions.

The storage of medical information increases in an astonishing rate. Most of specialty database will face the obstacle of executing efficiency. The system use the off-peak time to generate time-consuming tables and reports. That results some compromise of timeliness but make balance between efficiency and accuracy.

HIS is a huge machine and lack of reflex, therefore, invalid data accumulation is seldom detected by information administrators. Such invalid data may not result in financial loss but great medical statistics errors. Specific feedback from professional staffs is the best method to eliminate such errors. However medical staffs are not familiar with queries or data structures, they can not find the way to give feedback or explore the errors. The system does detect significant invalid data, which is eliminated for data integrity and make feedback to information administrators.

Incorrect medical behavior may cause serious consequence. The system can detect such behaviors and present the change if some improvement protocol has been executed. Continue quality control can be achieved through this process.

The productions of the system are abundant and variable. Some physicians of other department may question about the exact effect on decision-support. The impact on the decision-support can be evaluated through questionnaire evaluation. However the questionnaire can not comprise all aspects of decision-support. What we can improve is expand aspects of questionnaire evaluation. Besides, the system can form specific predict model[6] which is derived from answer of the specific question. The result of executing efficiency revealed a particular problem. The multiple queries executing time may approach one minute, which reflects necessity of SQL queries reformation and unsuitability of Access® for such application.

4.2. Future works

As taxonomies of decision-support system, the system can be classified as a data-driven decision-support system. For architecture[7] of decision-support system, we lack a model component, which should be created through answering specific questions belonged to our knowledge architecture. That will spend time to accumulate the final answers to form rules and expand the aspects of application. The image is very important for orthopedic surgeons and incorporation of image database will enhance the effect of decision-support. The next step is upgrading the system through creation of image database and extraction image data from Picture Archiving and Communication System (PACs).

4.3. Conclusion

This unique system for clinical and academic use of orthopedic departments, which possess functions of auto-extraction of patient data source from the hospital information system (HIS) and web interface for data queries. After a series of assessment and evaluation, the system do is able to play a positive role in the decision-support process.

Reference

[1] McManus, J. G.. "Orthopedic office medical records using COSTAR." Orthop Clin North Am 17(4): 581-90. (1986)

[1] Stoodley, M. A. and J. M. Sikorski. "OASYS: a computerized auditing system for orthopaedic surgery." Int J Biomed Comput 29(2): 119-31. (1991)

[1] Ruffin, M. "The importance of data warehouses for physician executives." Physician Exec 20(11): 45-7. (1994)

[1] Prather, J. C., D. F. Lobach, et al.. "Medical data mining: knowledge discovery in a clinical data warehouse." Proc AMIA Annu Fall Symp: 101-5. (1997)

[1] Kerkri, E. M., C. Quantin, et al.. "An approach for integrating heterogeneous information sources in a medical data warehouse." J Med Syst 25(3): 167-76. (2001)

[1] Lin S.H. Access the risk factors of transferring to intensive care units by artificial neural network in orthopedic practice. Annual meeting of orthopedic academy, Taiwan (2008)

[1] Power, D. J.. Decision support systems: concepts and resources for managers. Westport, Conn., Quorum Books. (2002)