Generic Data Modeling for Home Telemonitoring of Chronically Ill Patients

Jinsong Cai, Ph.D., Stephen Johnson, Ph.D., George Hripcsak, M.D.
Department of Medical Informatics, Columbia University, New York, New York 10032

ABSTRACT

Management of many types of chronic diseases such as diabetes and asthma relies heavily on patients' self-monitoring of their disease conditions. In recent years, internet-based home telemonitoring systems that allow transmission of patient data to a central database and offer immediate access to the data by the care providers have become available. However, these systems often work with only one or a few types of medical devices and thus are limited in the types of diseases they can monitor. For example, a system designed to collect spirometry data from asthmatic patients cannot be easily adapted to collect blood glucose data from diabetic patients. This is because different medical devices produce different types of data and the existing telemonitoring systems are generally built around a proprietary data schema specific for the device used. In this paper, we describe a generic data schema for a telemonitoring system that is applicable to different types of medical devices and different diseases, and show an implementation of the schema in a relational database suitable for a variety of telemonitoring activities.

INTRODUCTION

Chronic diseases such as diabetes, heart disease and asthma affect a significant number of people in the United States. Treatment of these diseases relies as much on patients' self-management as on the care provided by the healthcare professionals (1,2). Patient home self-monitoring using some type of data acquisition device has become an accepted practice in the management of these chronic diseases. The evolution of computer technology has led to an increasing number of computer applications that aim to assist patients in the management of these diseases. For example, a number of PC-based diabetes management software products have been used to download patient data from blood glucose meters and record patient self-management activities such as medications, diet and symptoms. More recently, home-based telemonitoring systems that connect patients and care providers through the internet have been developed and tested in a few academic medical centers (3,4). Commercial telemonitoring services such as data centers that collect data from asthma patients have also become available. At New York Presbyterian Hospital (NYPH), we previously developed an asthma home monitoring system that transmits patient data from a palmtop computer in patients' homes to the NYPH clinical information system and makes the data immediately available for review by their care providers using a web browser (3). This system also contains intelligent decision support tools that can alert physicians and patients when abnormal readings are detected.

Telemonitoring systems allow physicians to assess their patients' condition on a frequent basis and provide a new means of communication between the patients and their care providers. However a major limitation to the widespread use of these systems is that they are generally designed to work with only one type of chronic disease involving one or a few medical devices. This is because these systems rely on a central patient database that is optimized to one type of monitoring data. In general, different types of medical devices, and in many cases the same type of device from different vendors, produce data of different types. Furthermore, different system developers also tend to use their own document formats for data interchange. As a result, the software components that are built around the data become specific to the data from only one type of monitoring activity. A hospital that needs to monitor different types of diseases may have to build separate monitoring systems for each disease or medical device involved. Storing data in different data schema also complicates the task of integrating information systems in a healthcare environment. The heterogeneity of data schema used in different monitoring activities make it difficult to aggregate data across disease boundaries.

As a prelude to building a telemonitoring system applicable to different types of medical devices in the management of different chronic diseases, we concentrated on the creation of a generic data schema that can accommodate various types of data. In this paper, we survey patient data from five different monitoring devices and describe a generic data schema that is derived from these data. The feasibility of this data model is validated by an implementation in a relational database.

METHODS

Patient monitoring devices

In order to obtain a generic data schema for a wide range of medical data acquisition devices, we collected sample data from five different types of medical devices used for home care of chronically ill patients. In general, these devices are accompanied by software programs that download data from the devices to desktop or handheld computers and provide analyses and interpretations of the data. These devices include a spirometer (device and software from QRS, Inc.) for asthma monitoring, a blood glucose meter (device from Johnson & Johnson, software from Diabetic Meter Utility) for diabetes monitoring, a pulse oximeter (device from Datex-Omeha and software from ProFox) to monitor blood oxygen saturation, a blood pressure monitor (device from Pulse Metric), and a Holter monitor (device and software from Forest Medical). Although this is by no means an exhaustive listing of all medical devices, the principle used in this paper to obtain a generic data model can be generalized to other types of devices.

Generic data modeling

A detailed data schema was developed for patient data produced in these monitoring devices. This included a listing of all concepts and relations used to describe different monitoring activities. The detailed schema was then transformed into a generic schema that contained only a small number of generic concepts related to each other by one-to-many relationships as described by Johnson (5). While there were a number of ways to implement this data schema in a relational database, we used a simplified implementation in which each generic concept was implemented as a database table and the one-to-many relationships were represented using the foreign key constraints. The database was populated with sample data from the five medical devices.

RESULTS

A generic data schema was developed as shown in Figure 1. Only three general concepts – patient, monitoring_event, and monitoring_result – were found to be necessary in representing data in any given telemonitoring activity. The one-to-many relationships between these concepts indicate that each patient can have one or more monitoring events and each monitoring event can have one or more monitoring results. Furthermore, both the monitoring_event and the monitoring_result concepts allow nesting. One monitoring event may contain one or more sub-events, while one monitoring result may also contain one or more sub-results.

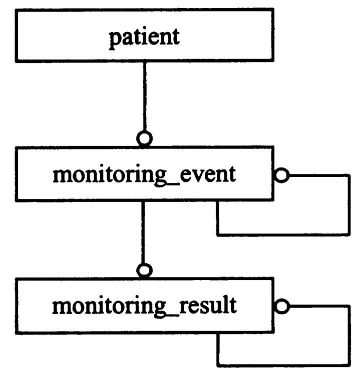

Figure 1. A generic data schema for telemonitoring. The symbol –O denotes a one-to-many relationship

The generic data schema could be implemented in a number of ways in a relational database. In this paper, we describe a simplified implementation in which each concept in the schema was implemented as a table in the database. The patient table contains all general information about the patient, including patient demographic, name of physician treating the patient, etc. The second table, monitoring_event, stores various types of events such as procedures or tests performed by patients during the monitoring period. The last table, monitoring_result, gives the detailed information about the procedure or test recorded in the monitoring_event table. Note that in a generic database schema, the various attributes that are represented as columns in a detailed database schema, become the values in a single column, and their properties are represented in the corresponding row in a different column. This design gives great flexibility to the database: the schema does not have to change when new attributes are added to the table.

Next we will use asthma and Holter monitoring as examples to illustrate how these database tables may be used to store data from different types of monitoring devices. In a typical asthma monitoring session, a patient performs up to three consecutive FVC tests. Each test records a flow volume loop and measurements of the forced vital capacity and other pulmonary indices. The monitoring_event table (shown in Table 1) indicates that a patient (identified by Patient_ID 1234567) performed three FVC tests on 11/22/95 in an asthma monitoring session. This asthma monitoring session contained three consecutive FVC tests. Note that all three tests have the same date-time in the group column, indicating that they all belong to the same session that is identified by this date-time value. The Patient_ID and Date-time form the primary key of this table; the Patient_ID is the foreign key that references the patient table (not shown).

Table 1. The *monitoring_event* Table in a Generic Database

Patient_ID	Date-time	Code	Group
1234567	11-22-95 8:35:30	[asthma monitoring session]	
1234567	11-22-95 8:35:37	[FVC test]	11-22-95 8:35:30
1234567	11-22-95 8:38:13	[FVC test]	11-22-95 8:35:30
1234567	11-22-95 8:42:21	[FVC test]	11-22-95 8:35:30
7654321	10-12-99 18:08:00	[Holter monitoring session]	

Table 2. The *monitoring_result* Table in a Generic Database

Patient_ID	Date-time	Seq#	Code	Parent	Num	Char	Coded value	bitstring
1234567	11-22-95 8:35:37	1	[FVC]		3.290			
1234567	11-22-95 8:35:37	2	[FEV1]		2.690			
1234567	11-22-95 8:35:37	3	[flow volume loop]					{FVC curve}
7654321	10-12-99 18:08:00	1	[ECG]					{ECG graph}
7654321	10-12-99 18:08:00	2	[heart rate graph]					{heart rate graph}
7654321	10-12-99 18:08:00	3	[hourly summary]					
7654321	10-12-99 18:08:00	4	[start time]	3		18:00		
7654321	10-12-99 18:08:00	5	[end time]	3		19:00		
7654321	10-12-99 18:08:00	6	[ventricular ectopics]	3	97			
7654321	10-12-99 18:08:00	7	[bigeminy]	3	2			

Table 3. Storing Data Elements of Different Monitoring Devices in a Generic Database

Device	# data elements produced by device	# data elements stored in tables		
		patient	monitoring_event	monitoring_result
Spirometer	38	4	3	31
Blood pressure	5	1	1	3
Blood glucose	8	1	1	6
Oximeter	516	6	1	509
Holter	440	12	1	427

The detailed information about each individual test is stored in the monitoring_result table (Table 2). The Patient_ID, Date-time, and Sequence number (Seq#) form the primary key for this table; the Patient_ID and Date-time form the foreign key that references the monitoring_event table. This table allows storage of a variety of data types, including numbers, characters, coded values and binary bitstrings. (However, in a real database implementation, it may be better to split the monitoring_result table into several tables, each optimized for one data type.) The measurements of the parameters of an FVC test (FVC and FEV1) are stored as numbers, while the flow volume loop is stored as image bitstrings. These parameters and the image belong to the same test because they all share the same Paient_ID and Date-time, the foreign key that references the monitoring_event table.

The same database tables can also be used to store Holter monitoring data. A Holter monitor records a continuous ECG that can be downloaded to a desktop computer and analyzed by a software program, The software program can generate values for a number of important parameters useful in the assessment of patient conditions. Using the same database tables, we can save all data produced in a Holter monitoring event. For example, the ECG graph, heart rate graph, and other computer-generated graphs are stored in the monitoring_result table as image bitstrings. The hourly summary is further divided into results for each individual hour. The numbers of ventricular ectopics and bigeminy episodes during the hour from 18:00 to 19:00 are saved as numbers and characters, as shown in the last four rows in this table (Table 2).

In order to find out whether this database design is sufficient in handling patient data in a wide range of monitoring activities, we tested the database using sample data collected from all five different medical devices. The devices were found to produce data elements of very diverse types, including free-text, numbers, coded data, graphs, and images. The database was capable of storing all data elements of all data types for the five different devices in one of the three generic database tables as shown in Table 3, validating the feasibility of the database design.

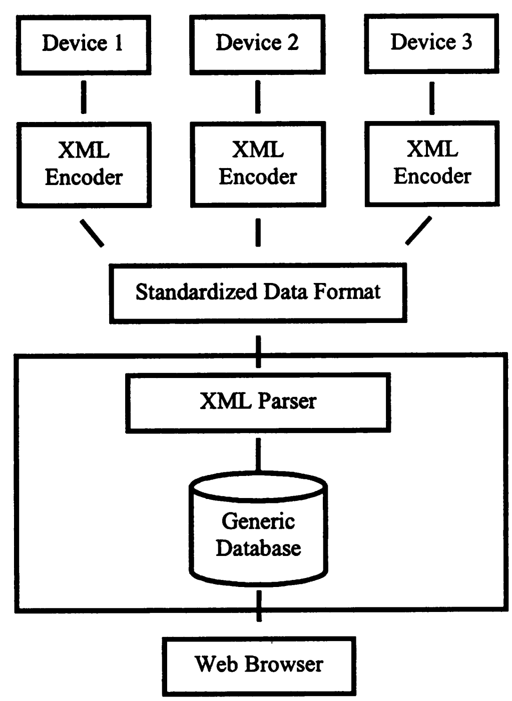

Figure 2. Information flow in a generic tele-monitoring system.

DISCUSSION

Home telemonitoring as a field has witnessed considerable growth in recent years. However, no generic data schema for a general-purpose telemonitoring system has been reported. We present here our initial step in building a generic telemonitoring system – the creation and validation of a generic data schema suitable for several different monitoring activities.

Using generic patient database to deal with the diversity and complexity of healthcare data has a long history in the evolution of clinical information systems (5,6). This not only provides a solid theoretic foundation, but also the successful experience that can be transferred to new medical domains. In general, a generic database is often used together with a data dictionary or a controlled vocabulary that provides definition and semantic knowledge of the coded data in the database. Generic data modeling has been successfully employed in creation of the clinical data repository at NYPH, the database that integrates various ancillary systems and contains clinical information for 1.3 million patients. HL7, the most widely used healthcare industry standard for data exchange in the United States, also employs a generic schema in its Reference Information Model (7).

Though not shown in this paper, it is also possible to create a standardized data interchange format for a generic telemonitoring system using the data schema presented here. If data from different monitoring activities are represented in a uniform data interchange format, it is possible to build a generic telemonitoring system that only have to deal with data in a standardized manner and is no longer constrained by the devices it interfaces with. Figure 2 shows the information flow in a possible generic telemonitoring system in which the standardized data interchange format is represented in the Extensible Markup Language (XML). In this architecture, the XML encoder residing on the patient computer is responsible for transforming device-specific data into a general format that can be mapped to the generic database schema. After data are transmitted to the server, an XML parser is used to extract data from the document before they are saved in the central database. Using this design, the software components and the database in the central server will only have to deal with data in a uniform manner, making the system device-independent.

ACKNOWLEDGEMENTS

This work was supported by National Library Medicine grants LM05627 "Linking Knowledge-Based Systems to Clinical Databases" and LM07079 "Research Training Grant". The authors are grateful to Drs. Jen-Hsiang Chuang and Stephen Wang for their helpful discussions.

REFERENCES

1. National standards for diabetes self-management education programs. *Diabetes Care* 1995;18:100-116.

2. Li JTC, Home peak expiratory flow rate monitoring in patients with asthma. *Mayo Clinic Proceedings* 1995;70:649-656.

3. Finkelstein J, Hripcsak G, Cabrera M. Telematic system for monitoring of asthma severity in patients' homes. *Medinfo* 1998;9:272-276.

4. Gomez EJ, Del Pozo F, and Hernando ME. Telemedicine for diabetes care: the DIABTel approach toward diabetes telecare. *Medical Informatics* 1996;21:283-295.

5. Johnson, SB. Generic data modeling for clinical repositories. *JAMIA* 1996;3:328-339.

6. McDonald CJ, Blevins L, Tierney WM, Martin DK. The Regenstrief medical record. *MD computing* 1988;5:34-47.

7. HL7 Reference Information Model. 1999; (http://www.mcis.duke.edu/standards/HL7/data-model/HL7/modelpage.html).