Medical Information Systems: Characterization and Challenges

Jorge C. G. Ramirez, Lon A. Smith and Lynn L. Peterson
Computer Science and Engineering
The University of Texas at Arlington
Arlington, TX 76019

Abstract

This paper examines the characteristics and challenges presented by medical databases and medical information systems. It begins with a survey of medical databases/information systems. This is followed by a list of challenges for database management systems generated by the needs of these systems. It concludes with a look at some systems which address these challenges. In the context of this background information, the database community is asked to consider whether the results of database research are reaching those who are making day-to-day decisions regarding design and implementation of medical information systems.

1 Introduction

For the moment, assume you work in the academic computing center of a large medical school. A busy surgeon arrives wanting to investigate an idea which has occurred to him, a possible relationship between observations made in several of his recent patients. What do you use to pursue this investigation? A likely source might be some information which is found only in the patient charts, or if you are lucky, data might be found in machine-readable form in an existing database. As a "quick-and-dirty" exploration, the form of the data is not terribly important, as long as it can be used to support the needed analysis.

However, in any long-term investigation, a more reasoned approach to data collection and analysis will involve a medical database or even a full-blown medical information system (MedIS). In deciding what form such a MedIS should take, a myriad of factors need to be considered, some theoretical and some quite practical. From a theoretical standpoint, issues of database evolution control, temporal/intelligent reasoning, and scalability are relevant. Practical issues include handling of multi-media data, data integration, mobile access to data, and probably most importantly, cost. As a final consideration, the system must be acceptable to and accepted by the users. Therefore, it must be easy to use (with truly user-friendly interfaces) and it must provide the users with what they perceive as measurable payback for their effort in using the system.

Medical literature is helpful in providing examples of how others handle similar design problems, that is, how the designers of a MedIS can likely proceed with its development. A reasonable question to ask is whether those in the database community will make the same design decisions as our hypothetical designer. If not, why not? A negative response might be attributed to their access to up-to-date database research. In this article, we would like to challenge the database community to consider whether or not database research results are finding their way to those who are making practical, day-to-day decisions regarding the design of MedISs.

In what follows, we investigate medical databases in an effort to characterize those that are predominant in the medical literature today. Next, we will propose some challenges for database management in this domain, and finally we will examine examples from various sources that exemplify attempts to address some of these challenges.

2 Characterization of Medical Databases

We will use the term MedIS to mean a system which incorporates one or more medical databases. But what is, then, a medical database? It seems clear that the term does not have a precise definition. At one level, it may merely mean a database that is involved, in some way, in the delivery of health care. This includes medical journal databases which are also called bibliographic databases, clinical information systems for patient care as well as billing, and medical research databases. Many people mean only the bibliographic when they say "medical database." A literature search on the term certainly supports this point of view. But to the medical information community, issues involving clinical information systems and research databases are at least equally important and dominate in any discussion of the term "medical database."

2.1 Bibliographic Systems

Perhaps the most widely known example of a bibliographic database is MEDLINE, though many others ex-

ist. Yale University's Current Contents [29] is one such system. Current Contents, which contains citations only, is updated weekly and is provided on-line. This can make the information available up to six months sooner than is possible through MEDLINE. Multiple bibliographic databases can be accessed through the Swedish system, Multi-Link. This system provides access to a number of bibliographic databases through a central system [45]. When viewing the investment of resources, it is thus apparent that there is a demand for the availability of such bibliographic systems.

For those interested in medical databases, however, bibliographic databases are important for another reason: they reinforce the role of standards. MEDLINE provides a standard set of descriptive elements for citations and abstracts of published literature, MeSH (Medical Subject Headings) and UMLS (Unified Medical Language System) provide standards for subject content, and MARC (MAchine Readable Cataloguing) provides a data exchange format. The significance of a standard vocabulary across medical databases cannot be underestimated. Components of the UMLS, currently in various stages of development, are: Meta-1 (its metathesaurus which derives base vocabulary input from MeSH, DSM-III, COSTAR, SNOMED, ICD, CPT), Semantic Network (linking concepts by relationships), Information Sources Map, plus access tools such as a knowledge-base user interface and a "coach" to assist in its use [5]. This standards effort should be examined before any individual efforts at vocabulary building are undertaken. Without evolving standard vocabularies, the development of multiple distributed bibliographic database systems will be impossible.

2.2 Issues Characterizing Medical Databases

As we searched the literature to determine what characterizes a medical database, it became clear that there are at least seven major recurring issues. Each of these is briefly discussed here.

2.2.1 Types of Data

Stead & Hammond [36] summarize the types of data that can be found in medical records systems. They list seven types: 1) Demographic, identifying and socioeconomic data; 2) General/non-time-dependent data, such as research protocols, tumor registry information and special patient needs; 3) Problems or diagnoses, including linkages to indicate the problem's origin, whether that problem is active or inactive, and dates of associated treatments; 4) Time-oriented (not resulting from an order), patient's subjective complaints and physical findings, which may include assessment of disease progress; 5) Data resulting from an order, lab tests and x-rays; 6) Therapies, including diets, immunizations, prescriptions, etc. and 7) Encounter/admission data, including dates, places, and provider of record for encounters.

We have discovered that there seems to be some debate about exactly what data is necessary to be recorded; however, the greater issue seems to be the level of abstraction of the recorded data. Consider for example storage of text and image data.

Free Text vs. Coding

The implementations of data storage run the gamut from all text to all coding, with many variations between them. Dorda, et al. [13] describe WAREL, a system that uses processed text and a retrieval algorithm which uses thesauri. Clark & Shea [8] present a free text system that is used for smaller database systems on PC-DOS based machines. GPSYCH, as described by Aisen & Lawlor [1], which is a combination of coded data and free text, is typical of many systems. The issue, identified by Whiting-O'Keefe, et al. [44], is structure vs. utility/acceptance. The more coded the data, the more structured the data, the higher its utility, but the lower its acceptance. A secondary issue involves the use of the data. When data is coded, it is usually intended for a specific purpose. Its availability for reuse in another capacity may then be limited by the initial encoding scheme.

Image Data

The literature has dealt more recently, but less well, with the storage of image data. Dayhoff, et al. [10] give a brief introduction to the Veterans Administration (VA) DHCP Imaging System which associates multimedia results with given medical procedures. Lowe [22] describes the Image Engine for retrieval of multimedia data including image, video and text through a MacIntosh interface. However, Zink & Jaffe [47] meet the entire issue of image data head-on. They point out that not only is the analysis and archiving of images for any single patient not well addressed, but there is little attempt to gather data from images of different patients for the purpose of obtaining relevant disease knowledge. According to Dayhoff, et al. [12], integration of all of a patient's data, including clinical history, pathology, laboratory work-up, image data, results of other relevant tests, as well as simultaneous access to relevant medical literature and other reference material, is perceived as the single greatest barrier to the use of the full potential of computer networks.

2.2.2 Types of Records

Our survey reveals a wide range of purposes for medical data records: some are clinical, many are managerial, and others are not so obvious. Individual patient care and population data for research are obvious reasons for utilizing a MedIS. Likewise, clinical audits and health

service management deal with managing care and costs. Maybe not so obvious, however, is use of records for decision support and legal documentation.

The most influential work in the past thirty years, which addresses the issue of medical record maintenance, is Weed's Problem Oriented Medical Record[43]. The goal of this approach is to facilitate computerization, a goal which seems not to have been accomplished well. The fundamental entity of this record is a "problem." A problem list gives the state of the patient and all data is associated with a specific problem. Rector, et al. [31] find that using a problem as a fundamental entity is too cumbersome. They argue that an observation should be the most fundamental entity. They further state that these observations should be recorded at whatever level of abstraction the caregiver perceives them, not only symptoms, signs and abnormalities as prescribed by Weed. Two major types of records, event-based and patient-based, have descended from Weed's approach. Most systems allow for retrieval by either approach, though one or the other is the actual basis of the system.

Event-Based Records

No matter what the goal of the system, and they do vary, the idea of storing records in the system in chronological order, as they occur, is the basis for many systems. A couple of the most notable are the Summary Time-Oriented Record (STOR) [44] and HELP-LDS [30]. Though these systems are mature, they still well exemplify some key issues in the use of event-based records.

The STOR system addresses the issue of unstructured vs. coding, but fails to address the issue of longitudinal records for individual patients. The goal of this system is to provide on-line, real-time display access of clinical information. In order to reach this goal, a centralized summary database of clinical patient information has been developed. To address storage problems and access time, coding has been instituted.

The HELP-LDS system developers have analyzed data use and have found that data output far exceeds data input. Furthermore, the most frequent access is at the event level. Given that the goal of the system is decision support, their conclusion was to go event-based. They also chose to use to coded data, which is ironic given their stated fundamental belief in multiple use of data.

Patient-Based Records

Most systems that use the patient-based record cite patient care as the first priority in developing a computerized system. Significant systems in this area include The Medical Record (TMR) [36]. This system evolved as a result of user requests over time, not as a result of an original comprehensive design. The authors emphasize storage of data according to its true meaning, not in any interpreted form.

PEN & PAD [32] also places patient care as the priority with emphasis on a faithful recording of what clinicians have heard, seen, thought and done. The authors divide observations into two levels: 1) Direct observations - actual observations recorded by the observer, and 2) Meta-observations - transcribed data entry and decisions/synthesized observations. Heathfield, et al. [18] describe an extension to this system which gives a basis for collaboration between various types of clinicians. Each is given a view of the data that is most appropriate to the specific discipline.

We would be remiss if we did not also note contributions of the Regenstrief Medical Record System [25]. The major goal of the system is to provide a summary computer stored records, not unlike the STOR system. The integration of electronic data capture with manual data capture, as well as coding information for long-term storage vs. retaining full text for short term storage, is an approach used here that is a unique combination in this well-known system.

2.2.3 Database Organization

Relational database organization is in frequent use in medical systems, and especially those with a specialized emphasis, such as results review and decision support in the Columbia-Presbyterian system [20]. In [16], the choice of a relational organization is based on theoretical and practical considerations and allows flexibility in choice of the type of record (POMR or conventional). But the object-oriented database organization, deductive databases, and temporal database concepts have also found practical application in medical systems.

Object-oriented databases allow development of information systems that focus on the modeling of the data to be more closely representative of real world entities. By the use of encapsulation, this design paradigm allows systems to be developed at an abstract level, thus leaving the details to be developed after a good high-level understanding has been developed [19]. However, the disadvantages of the actual implementation and testing of this type of system can outweigh the benefits of an easier-to-understand design capability. It is argued, however, that the modeling of data in ways that more closely represent the real world reduces design errors and promotes consistency in the system.

Deductive databases are successfully used in both clinical and research settings. They have the benefit of being based on standard relational database systems and the power of self-defining new relations [33]. These new relations are built based on user-supplied rules that "deduce" information that can not otherwise be easily observed from the data. While these types of systems are good at delivering results that can be observed with

some effort, they are only a primitive approach to learning all that can be learned from the existing data.

The use of temporal database design concepts adds an extra dimension to data analysis. Most data is viewed as one dimensional as it is collected. However, in medicine it is the observation of the course of disease over time that is crucial. Therefore, the addition of the temporal aspect adds another dimension to the collected data. Temporal abstraction mechanisms, such as described by Shahar, et al. [34], provide a first approach to automating this type of observation. Temporal data operators, as described by Das, et al. [9], provide further capabilities at analyzing data over time. These approaches, combined with TimeLineSQL [9] for querying these data, offer a strong foundation for further advancement in automating observation of the course of disease.

Johnson, et al., [20] discuss the notion that an integrated view of all data in a system can be provided by a conceptual data model. This model also bridges the gap between the data organization as viewed by the DBMS, the physical data model, and as viewed by the individual user, the logical data model. In relational databases, the relational schema can often play the role of conceptual layer. But a higher level of abstraction for data modeling can be provided by entity-relation diagramming, knowledge bases of logical rules, and object-oriented analysis. Significantly, the application of any one of these techniques can allow the same model to be used with different DBMSs.

Proposals for a standard data model of clinical information are emerging, e.g., the model proposed for use in the HL7 (Health Level Seven) [35] and MEDIX standards for clinical information exchange. Both of these standards are intended to facilitate the electronic transfer of patient data, both specifying messages, record types and data elements. HL7 (Health Level Seven) is a health-care specific, application level protocol in use at this time, while MEDIX is not yet fully developed. Some of the entities of the MEDIX model include: location, such as offices or beds; person, including patient and health care professional; organization, for example family or provider organization; as well as entities comprising the actual patient chart and entities needed for financial management. Standards for clinical data items are evolving, as described in the UMLS efforts mentioned above.

2.2.4 Mass Data Storage

The standard way to handle the archiving of patient data uses tape storage and/or demountable disk packs for data on patients discharged, say, 6 months ago. Archiving can also be managed using a virtual database, in which a user application interacts with a virtual access mechanism requesting retrieval of information from virtual files, which either must be physically resident on the central computer or retrieved from physical files on a database server [37]. CD-ROM is a reasonable medium for static information, e.g., bibliographic and historical patient data. New technology for massive data storage, such as E-Mass, E-Systems' 10,000 terabyte mass storage device, are becoming available. This type of device will greatly impact actual physical storage and address longitudinal patient care, as well as researchability issues.

2.2.5 Distributed Data Issues

The typical heterogeneous computing environment found in most health care institutions provides a significant impediment to the development of an electronic medical record. Many institutions have built interfaces to major systems (e.g., the clinical lab system, the admission system), but these "one of a kind" interfaces are difficult to write and more difficult to maintain or replicate. Therefore, much data is not combined into a comprehensive computer-stored record, and paper remains the predominant medium for the patient's medical record [35]. However, once the issue of electronic transfer of patient data is addressed (e.g., with the HL7 protocol), desires increase and users wish to be able to access and use data stored throughout such a heterogeneous computing environment.

Whatever the actual method of distributing data or databases, the goal of distributed data systems is to make the user think he or she is using a single database. Heterogeneous distributed systems, representing the most common type of distributed system, involve multiple databases, usually on multiple computers, each with its own local schema to be completely or partially integrated to form a unified schema through which retrieval and/or update is possible. Marrs, et al. [24] describe access to clinical data distributed among multiple satellite information systems, and the specific schema and data integration techniques used in their application. They note that, not surprisingly, the most laborious step in schema integration is that of discerning the intended or implied semantics of legacy data systems. Without a detailed knowledge of the semantics of the data, it is difficult to recognize similar or related concepts (e.g., are account number and registration number related?).

Decentralized databases, on the other hand, are characterized by an integrated (but physically separated) database manipulated by a standard set of tools. The VA's DHCP [11] is the best example of a decentralized system, with separate databases at separate sites being managed by the same software. This type of system addresses the need for standardization in medical data management while also addressing cost concerns. The VA situation is such that having a national medical database network that reaches their approximately 200

sites is cost prohibitive. The solution is the decentralized approach.

2.2.6 Interface Issues

Good interface design criteria dictate that the interface should be: easy to use, easy to access, have a rapid response time, and protect confidentiality [32]. Recently, expectations for a reasonable user interface have risen to the point where a graphical user interface (GUI) is more the rule than the exception in new medical applications. When used in data collection, a user interface which provides more feedback of information to the caregiver/user will be more readily accepted and will tend to encourage better collection of data [17]. As shown by Cimino & Barnett [6], a display of clickable nodes and arcs can help a user navigate through a network of options in selecting information for retrieval. Finally, hypertext is a very effective user interface to medical information. The difficulty in preparing information for subsequent access through hypertext (i.e., creating the links between concepts in the text) can be expedited, as shown by Cimino [7], through the use of a medical concept space with concepts linked hierarchically, to which sections of text can then be linked. The effect is that a user can browse the text using the links of the concept space without requiring a hand-creation of links in the text.

2.2.7 Decision Support Issues

The way in which knowledge is to be integrated with a database or a system of databases is the primary decision support issue. In deductive databases, obviously the decision is to imbed the knowledge in the database itself. But, other options exist. Independently built and maintained knowledge bases can exist and interact with separate databases, all controlled by a workstation [39]. As another alternative, the user can interact with an expert system which has access to retrieval databases, reasoning databases, and separate knowledge bases [2].

Such decision support systems are conceptually feasible, but usually require custom coding. Nguyen and Marguiles [28] present the Clinical Event Monitor. It is designed to review classes of clinical events, such as readmission of an emergency patient within 96 hours, using a virtual relational database combined with a rule base. Schwaiger, et al. [33] describe NIMON, which is designed to aid in the analysis of renal function data. NIMON utilizes deductive database technology and a knowledge base.

The most comprehensive approach to the development of decision support systems can be found in the efforts of Tuttle, et al. [40]. They are attempting to develop a set of public domain standards for a system they call Knowledge-Server. It is designed to be an intermediary between the caregiver's knowledge needs and the knowledge sources that are available and relevant.

3 Challenges for a Medical Database Management System

In considering the characteristics of MedISs just discussed, one might choose to view them as requirements which can easily be placed on a system being developed and to believe that a system which can handle all of these needs might exist and might even be widely available. Examining this notion raises several questions.

First, how have "all of these needs" been identified? Though it is easy to theorize, a comprehensive examination of the true needs of a health-care delivery/management/research infrastructure should speak to that issue. Many of the systems described above are systems that have been designed piece-meal, i.e., as needs developed, a new component has been designed. These systems have also been developed piece-meal, i.e., as money allowed, new components were developed. However, it is clear that the needs as they exist for the systems mentioned above have evolved over time and are fairly comprehensive. Since none of the referenced systems have been developed with a comprehensive design, many desirable features are not included.

Second, are there commercial products (or other maintained software) which are available to provide support for a variety of data formats, including image data and free text data, in addition to standard types of data? Some products can support all of these types of data at once. However, the availability of these products still does not address many of the needs of comprehensive MedISs.

Finally, what actually are "all of these needs?" What requirements should we be able to place on software that is needed to support MedISs? The following is a list of seven significant needs that have become apparent as we have surveyed the medical database literature.

Database Evolution Control

One of those needs is flexibility of structure of the database itself, known as database evolution control. As the types of data collected change over time, the structure needed to support that data collection must also change. This capability might be required after a merger of medical organizations that have possibly very different database organizations, but need to create a system that functions like a single MedIS. The advent of object-oriented databases have brought a type of technology that might be utilized to address this need.

Data Integration

Data and records associated with a specific patient can be located in a variety of databases/locations. This need might arise when dealing with patients that live part of the year in one state and part in another or patients who have recently moved, or in working with

patient records some of which may be at the physicians office and others at a hospital/medical center, or with patient records stored in different places within a hospital/medical center. One roadblock to success in this area is the significant number of doctor's offices, clinics, and even hospital/medical centers which do not maintain patient data (other than billing information) on-line. The majority of on-line systems deal specifically with billing, insurance and accounting and not with patient medical records. Without a data consolidation and presentation mechanism bringing together all of this data, it is not possible to present an integrated patient picture to physicians, an integrated facility picture for administrators, or an integrated data picture for researchers.

Scalability

This is the ability of a system to grow to meet its demands. If a system is scalable, then as demand increases the system can be modified to handle that need. One dimension addresses the issue of an ever-increasing volume of data. As patients grow older and more data is generated for the average patient, the crucial question is how to determine what data is archivable, i.e., data which has aged sufficiently that it is no longer significant for medical care purposes.

A second dimension of scalability is from a research perspective. Longitudinal access to patient data is always desirable. Therefore, the development of new and inexpensive mass storage devices is essential to the success of long-term, large-scale MedISs.

But developments beyond hardware are also required. New techniques for addressability of large amounts of data may be needed by a DBMS. Issues also involve access to the data, such as going from one receptionist entering data to a hundred clerks entering data. From a user perspective, the issue of response time is also key. As the system grows, response times can degrade, resulting in unhappy and frustrated users.

Temporal Data Issues

Management of temporal data poses its own set of difficulties for database systems not specifically designed to address those issues. Some types of temporal analysis can be provided by imbedding intelligence in the database to detect typical trends in patient data over time. As unexpected medical situations develop, a mechanism for doing temporal analysis of trends in patient data can be helpful.

Intelligence Issues

Incorporation of intelligence can take one of several forms. Intelligence can be added by providing access to external knowledge or information sources such as expert systems or bibliographic sources. Alternatively, it might mean incorporation of knowledge base technology. Such technology can be used to assist physicians by providing alerts or warnings regarding entered data that is outside specified norms or combinations of data that indicate serious problems or delicate/unstable situations.

Knowledge base technology can be augmented by using aspects of "active databases" for the incorporation of both temporal and intelligent information. Example: a database has a temporal value such as a patient's temperature. After a length of time, the value in the database is no longer valid and the system should try to get that value updated if it is being used in decision making. The "active database" portion of the system recognizes this and does something about the invalid temperature. The intelligence is shown in that this is not just a time-dependent trigger, but there are rules that have been put in place to determine when a temperature is invalid.

Mobile Data Access

Mobile access to data is emerging as an important need. From an emergency point-of-view, vital information can be provided regarding an accident/disaster victim if the patient is identifiable and the data is accessible. From a social perspective, with a significant number of people homeless, the delivery of health care to this segment of the population and maintenance of the appropriate records and data will require mobile access.

User Acceptance

Last here, but certainly not least, user acceptance, i.e., availability of truly user-friendly interfaces, is an old, but extremely important issue. Any system that reaches to the physician office or clinic level must be intuitive and provide the medical staff with some real payback in their own environment. They will not see the point of wasting their time on working with an inefficient, cumbersome system for the benefit of some researcher.

4 What is Currently Available to Meet These Challenges?

Systems in a variety of domains have been developed which meet some of these challenges. They include systems that come from both the commercial sector and from academia. In the area of accounting systems, the database has become an indispensable tool. Systems such as Real World and Peach Tree, standard norms for off-the-shelf accounting systems, have features including: automatic classification, patterned statement generation, and multiple source analysis. These products utilize user-friendly features such as pull-down menus and hot-keys which make them easy to use. Yet, even though these packages are used as part of medical database systems (Real World has an additional medical package), they do not attempt to do anything but

accounting. For instance, the medical package offered as part of the Real World product does provide additional terms and classification in a medical vocabulary, but does not provide even the simplest capabilities for tasks such as patient record storage.

Likewise, MedIS can benefit from considering systems which are seen in the area of library sciences, specifically bibliographic databases. Some of the features present in more general systems, like INMAGIC [42], should be captured. These features include: menus, pop-up windows, and fill-in-the-blank searching. Generally, on-line, full-text searching techniques have been used successfully on systems like DIALOG, BRS and NEXIS [38] which can aide the researcher and the physician with literature searches. BRS, in particular, has focused on the medical market, providing listings of numerous journals for medical research. Yet, these systems only work within a single domain and do not provide the capabilities which are desired in a complete MedIS.

Bibliographic databases have also found a home on CD-ROM. There are many products, such as PsycLIT and SearchMagic [27]. These products provide an inexpensive collection of literature with elaborate GUI, thus making searching both easy and user friendly. Update is slow as compared to on-line services and hence is a drawback for these systems. But, with the CD-ROM technology, there have been attempts to provide for more than just the single domain in medicine. A case in point is Micromedex, which has a "computerized clinical information system (CCIS)" [15] that includes encyclopedic-type knowledge in areas such as toxicology, pharmacology, and emergency medicine. Also, the system furnishes the capability to capture patient information, both domain-based and free-text, for the creation of on-line charts. Yet, CCIS falls short of being a complete MedIS because of a lack of capability in areas such as accounting and billing. Consequently, in a medical environment, each department maintains its own data and is not efficiently integrated so that this information can be utilized by all accessing departments.

In addition to the commercially available products, the academic world has done much to address some of the challenge issues. A conceptual approach for a data model which embraces clinical data, operation management and strategic information about care evaluation, cost-effectiveness and planning is being developed by the INFORM project in Europe [44]. This project is a collaboration between institutions in 5 countries. The objectives are to improve the quality, efficiency and cost-effectiveness of patient care. The three conceptual data groupings are supported by reference data, such as standard protocols, drug information, etc.

The issue of temporal aspects of data is addressed by a project at the University of Ulster, Northern Ireland [21]. The authors present a new architecture for the handling of time-varying data, which also addresses the issue of data integration, through a heterogeneous distributed capacity. The system is similar to a multidatabase approach for distributed systems, but the time-varying attributes are stored, processed and manipulated automatically by the temporal system itself. The system is broken into four modules: Global Temporal Module, Historical Management Module, Current Management Module and Future Management Module. Queries are posed through an extended SQL language, similar to TimeLineSQL. Other issues to be addressed include processing strategies in the distributed environment and dynamic user interfaces where results can be displayed in different formats, orientations and styles.

Work is being done in Boston to address the need for mobile access to data. Chueh & Barnett [4] describe a client-server, distributed database strategy with dial-in capability. This system has been developed for Boston's Healthcare for the Homeless Program. It utilizes portable personal computers which contain reference databases and provides modem access to a central database which contains patient data. This methodology reduces the amount of data that needs to be transmitted over the phone connection.

In Italy, work is being done on a system that addresses the issues of data integration and user acceptance [3]. The goal is to design an efficient distributed computer system to link the primary care community with general and specialty hospital departments (secondary care organizations). Additional requirements are that it should not hinder medical care, it should focus on the data needed during clinical examinations, and it should allow for the needs of the various users with different roles in a health care system. This system is based on a Minimum Basic Data Set for Ambulatory Care (MBDS-AC) and claims that the design and implementation of additional modules for the system should be an easy task, implying an open system architecture.

At University College, London, the applicability of object-oriented methodologies is being studied [14]. A significant factor in the rationale is that the application and underlying data model lends itself to an object-oriented approach. The authors point out that "In modelling medical information, the major difficulties arise because this information is multidimensional, has many viewpoints, large amounts of data need to be processed and also the relations between various data items may be very complex." [14, p. 755] The issue of user-friendliness is also addressed by the contention that "to become more widely used, it will be necessary to make data entry, data coding, and data retrieval much more efficient by providing a good human computer interface (HCI) and by incorporating facilities for coping with synonyms and ambiguities" [14, p. 754].

At the University of Giessen, Germany, the issues of data integration and intelligence are addressed through modifications to a HELP-LDS kernel [26]. The goal is the integration of decision monitoring as a type of med-

ical decision support. Decision monitoring is defined as typically sending alert messages to predefined destinations. This project uses a relational DBMS to support SQL, and requires domination of structured data over free text input. The system is organized as groups of cooperating processes which communicate in a client-server mode, so that no central component can make a departmental system unable to run in stand-alone mode.

One of the more comprehensive attempts is being made at University Hospital of Brussels (AZ-VUB) [39]. This project addresses the issues of flexibility, data integration, and scalability. The authors classify their project as an attempt to develop and implement a 2nd generation of hospital information system (HIS). The original, or 1st generation, system was focused on automating administrative and managerial tasks. The focus of the 2nd generation system is to meet the primary requisites for a medical record, which they define as "the registration of a patient's clinical and administrative information from various departments in order to build a longitudinal medical record that ties outpatient and inpatient clinical information in a single patient clinical record." [39, p. 52] This is accomplished through an open software architecture base on the UNIX operating system, and a two-level database architecture that emphasizes long-term vs. short-term data and includes an emphasis on the digitalization of diverse pictorial information.

Another system that incorporates the object-oriented approach, while addressing the issues of flexibility, data integration, and user acceptance, is the SYGMA system at Graz University Hospital, Austria [23]. Features of the system include: patient admission, patient management, and recording of examination-specific parameters. The system provides patient data and management information while allowing for easy development of add-on applications.

It seems that, quite properly, systems are being developed to meet specific individual needs, and that the need for a database system to be "all things to all people" has not yet arisen. For this reason, we suspect, little effort has yet been invested in the development of systems that try to capture a preponderance of the features which seem desirable.

5 Conclusions

We have presented a look at medical databases in an effort to characterize those that predominate in the medical literature today, and have looked at some challenges for database management in this domain. We have looked at existing systems which address those challenges, and have observed that most of these systems do so selectively.

We might speculate on why we do not currently have a fully-developed integrated DBMS package. Several possibilities come to mind: 1) There is a lack of domain experts who are conversant with all aspects of MedIS. 2) The desire to work with existing software as opposed to a new integrated system which might better suit the needs of the organization is strong among component designers of MedIS. 3) The obvious economic investment in systems of this magnitude and size discourages their development. 4) There exists a fear of creating a system which attempts to accomplish everything yet in actuality only does a few things well. A direction toward the solution of these deficiencies may be to build a system around the concept of an open system architecture. It is here where each module can communicate with each other easily and modules can be created, modified, and removed without destruction of the entire system.

We are aware that changes can come rapidly, and that readers may wish to keep current in this area. Good sources of current MedIS information from the medical point of view include (among others) the yearly Proceedings of the Symposium on Computer Applications in Medical Care (SCAMC) and monthly and bimonthly journals such as Computers in Biomedical Research, M. D. Computing, Computer Programs in Biomedicine, and Methods of Information in Medicine. An abbreviated bibliography of relevant references is included in this paper, but a more complete one can be requested on-line using the email address: ramirez@cse.uta.edu.

6 Acknowledgements

The authors would like to thank Gretchen K. Cormier, Consultant for CW Systems, Inc., Austin, TX for substantive contributions to the direction of this article. Thanks also to Dr. Abdelsalam (Sumi) Helal, UT-Arlington, for guidance on this project from the start.

7 Bibliography

[1]. Aisen, P.S. and B.A. Lawlor. GPSYCH: Clinical Management Software for a Geropsychiatry Division. SCAMC '92 p 322.

[2]. Berman, L., M.R. Cullen and P.L. Miller. Automated integration of external databases: A knowledge-based approach to enhancing rule-based expert systems. Comp Biomed Res 26 p 230.

[3]. Bernabei, A., V. Curro, A. D'Atri and G. La Cava. A distributed system for the integrated management of general and subspecialist pediatric outpatient clinic. MEDINFO '91 pp 101-105.

[4]. Chueh, H.C. and G.O. Barnett. Client-Server, Distributed Database Strategies in a Healthcare Record System for a Homeless Population. SCAMC '93 pp 119-224.

[5]. Chute, C.G., Y. Yang, M.S. Tuttle, D.D. Sherertz, N.E. Olson, and M.S. Erlbaum. A Preliminary Evaluation of the UMLS Metathesauras for Patient Record Classification. SCAMC '90 p 161.

[6]. Cimino, C. and G.O. Barnett. Standardizing Access to Computer-Based Medical Resources. SCAMC '90 p 33.

[7]. Cimino, J.J. As We May Think: The Concept Space and Medical Hypertext. Comp Biomed Res 25 p 238.

[8]. Clark, A.S. and S. Shea. Free Text Databases in an IAIMS. SCAMC '91 p 224.

[9]. Das, A.K., S.W. Tu, G.P. Purcell and M.A. Musen. An Extended SQL for Temporal Data Management in Clinical Decision- Support Systems. SCAMC '92 p 128.

[10]. Dayhoff, R.E., D.L. Maloney, T.J. Kenney and R.D. Fletcher. Providing an Integrated Clinical Data View in a HIS that Manages Multimedia Data. SCAMC '91 p 501-505.

[11]. Dayhoff, R.E., D.L. Maloney, P.M. Kuzmak, A. Sadan and W. Majurski. Integrated Imaging Workstations using MS-DOS and UNIX/X Windows. SCAMC '91 pp 965-967.

[12]. Dayhoff, R.E., D.L. Maloney, P.M. Kuzmak and B. Shepard. Integrating medical images into hospital information systems. J Digital Imaging 1991, 4(2) pp 87-93.

[13]. Dorda, W., B. Haidl and P. Sachs. Processing Medical Natural Language by the System WAREL. Meth Inform Med 27 p 67.

[14]. Eliasz, A.W. The design and implementation, using an object oriented methodology, of a user friendly primary health care patient management system based on the ICPC classification. MEDINFO '91 pp 754-758.

[15]. Fishman, D.L. Computerized Clinical Information System-CCIS for MicroMEDEX. Database(Weston, Conn.) 1992 15(Apr) pp 58-62

[16]. Gouveia-Oliveira, A., and L. Lopes. Formal Representation of a Conceptual Data Model for the Patient-Based Medical Record. SCAMC '93 pp 466-470.

[17]. Heathfield, H.A and J. Kirby. PEN&PAD (Elderly Care): Designing a Patient Record System for Elderly Care. SCAMC '93 pp 129-133.

[18]. Heathfield, H., J. Kirby, A. Nowlan and A. Rector. PEN&PAD (Geriatrics): A Collaborative Patient Record Systems for the Shared Care of the Elderly. SCAMC '92 p 147.

[19]. Jean, F-C., T. Thelliez, J-J. Mascart and P. Degoulet. Object-oriented Information System in the HELIOS Medical Software Engineering Environment. SCAMC '92 p 595.

[20]. Johnson, S., C. Friedman, J.J. Cimino, T. Clark, G. Hripcsak and P.D. Clayton. Conceptual Data Model for a Central Patient Database. SCAMC '91 p 381.

[21]. Ling, D.H.O., D.A. Bell and I.R. Young. A Temporal Database Management for Medical and Health Care Systems. MEDINFO '91 pp 121-126.

[22]. Lowe, H.J. Image Engine: An Object-Oriented Multimedia Database for Storing, Retrieving and Sharing Medical Images and Text. SCAMC '93 pp 839-843.

[23]. Madjaric, M. and G. Gell. SYGMA: A general Object-oriented Medical Application Development System. MEDINFO '91 pp 127-131.

[24]. Marrs, K.A., S.A. Steib, C.A. Abrams and M.G. Kahn. Unifying Heterogeneous Distributed Clinical Data in a Relational Database. SCAMC '93 pp 644-648.

[25]. McDonald, C.J., B. Tierney, M. Overhage, D. Martin, B. Smith, C. Wodniak, L. Blevins, J. Warvel, J.M. Johnson, L. Lemmon, and T. Glazener. The Regenstrief Medical Record System - Experience with MD Order Entry and Community-wide Extensions. SCAMC '93 p 953.

[26]. Michel, A., P. Sebald and J. Dudeck. Some Aspects of the Implementation of Decision Monitoring Logic within a Heterogeneous Hospital Information System. MEDINFO '91 pp 925-929

[27]. Nahl-Jakobovits, D. and C. Tenopir. Databases Online and on CD-Rom: How Do They Differ, Let Us Count the Ways. Database(Weston, Conn.) 1992 15 pp 42-50.

[28]. Nguyen, L.T. and D.M. Margulies. The Design of a Rule-based Clinical Event Monitor in a Multi-Vendor Hospital Computing Environment. SCAMC '92 p 432.

[29]. Paton, J.A., A. Belanger, K-H. Cheung, S. Grajek, K.A. Branch, N. Ikeda, L. Sette, P.L. Miller and R.K Fryer. Online bibliographic information: Integration into an emerging IAIMS environment. SCAMC '92 p 605.

[30]. Pryor, T.A. The HELP Medical Record System. M D Computing 5(5) p 22.

[31]. Rector, L., W.A. Nowlin and S. Kay. Foundations for an Electronic Medical Record. Meth Inform Med 30 p 179.

[32]. Safran, C., D.M. Rind, R.M. Davis, J. Currier, D. Ives, D.Z. Sands, W.V. Slack, H. Makadon and D. Cotton. An Electronic Medical Record that Helps Care for Patients with HIV Infection. SCAMC '93 pp 224-228.

[33]. Schwaiger, H., M. Haller and U. Finsterer. A Framework for the Knowledge-Based Interpretation of Laboratory Data in Intensive Care Units Using Deductive Database Technology. SCAMC '92 p 13.

[34]. Shahar, Y., S.W. Tu and M.A. Musen. Temporal-Abstraction Mechanisms in Management of Clinical Protocols. SCAMC '91 p 629.

[35]. Sideli, R., S. Johnson, M. Wechler, A. Clark, R. Simpson and C. Chen. Adopting HL7 as a Standard for the Exchange of Clinical Text Reports. SCAMC '90 pp 226-229.

[36]. Stead, W.W. and W.E. Hammond. Computer-Based Medical Records: The Centerpiece of TMR. M D Computing 5(5) p 48.

[37]. Stevens, L.E., S.M. Huff and P.J. Haug. The Development of a Virtual Database to Provide On-Line Access to a Large Archive of Clinical Data. SCAMC '92 p 600.

[38]. Tenopir, C. and S.Bergland. Full-Text Searching on Major Supermarket Systems. Database(Weston, Conn.) 1993 16(Oct) pp 32-4.

[39]. Timmers, T., E.M. van Mulligen and F. van den Heuvel. Integration of an Object Knowledge Base into a Medical Workstation. SCAMC '91 p 654.

[40]. Tuttle, M.S., D.D. Sherertz, L.M. Fagan, R.W. Carlson, W.G. Cole, P.B. Schipma and S.J. Nelson. Toward an Interim Standard for Patient-Centered Knowledge-Access. SCAMC '93 pp 564-568.

[41]. Van de Velde, R. Overview of an Architectural Approach to the Development of the A.Z.-V.U.B. Distributed Clinical Information System. MEDINFO '91 pp 51-56.

[42]. Veccia, S.H. INMAGIC Plus for Libraries: It's a Library-in-a-Box!. Database(Weston, Conn.) 1993 16(Oct) pp 44-8+

[43]. Weed, L. L., Medical Records that Guide and Teach. N. Eng. J. Med. 278, p 593-600, 652-657.

[44]. Whiting-O'Keefe, Q.E., A. Whiting and J. Henke. The STOR Clinical Information System. M D Computing 5(5) p 8.

[45]. Wu, G., H. Ahlfeldt and O. Wigertz. MultiLink - An Intermediary System for Multi-Database Access. Meth Inform Med 32 p 82.

[46]. Yates, C.E., M.S. Leaning, D.L.H. Patterson, C. Ambrosos and S.T. Kalli. Data Modelling for Intensive Care within the INFORM Project. MEDINFO '91 pp 116-120.

[47]. Zink, S. and C.C. Jaffe. Medical Imaging Databases: A National Institutes of Health Workshop. Investigative Radiology 1993, 28(4) pp 366-372.