DESIGNING A DATABASE SYSTEM FOR THE DIVISION OF RHEUMATOLOGY
Margaret G. Peterson, Ph.D., Analyst/Programmer
Trudy J. Lerer, M.S., Biostatistician
Marcia A. Testa, M.P.H., Ph.D., Chief Biostatistician/Director
Biostatistics Research Center
University of Connecticut School of Medicine
Farmington, Connecticut 06032

Abstract
A computerized medical information system was designed for the Division of Rheumatology, University of Connecticut School of Medicine. Several issues pertaining to systems analysis, data collection instruments, database schema and database software will be discussed in relation to the design process. The proposed system contains a hierarchical database structure and allows for networking, data screen forms entry, and multi-level security control.

Introduction
Non-computerized medical records are extremely limiting to medical investigators and clinicians in their attempts to access, analyze and summarize medical information on a series of patients. While manually accessed medical charts and records have been the norm for most subspecialty medical information systems, recent advances in computerized database management systems offer new alternatives for solving problems of data accessing and processing. The example discussed here describes the existing information system of a division of rheumatology and outlines the basic steps toward designing a computerized medical records system.

Current System: Manually Accessed Medical Records
The Division of Rheumatology at the University of Connecticut has many functions including operation of patient clinics and wards, teaching and research. These functions result in a variety of types of information collected on numerous patients. The current patient database containing this information is recorded on paper in the form of medical charts and is physically stored in filing cabinets in the division’s office.

Over the years, thousands of records have been collected presenting the research worker with a daunting task when searching for any particular type of record. Much time is wasted retrieving records for the desired combination of symptoms or dates and random selection of particular records is extremely difficult if not impossible. Furthermore, all records more than ten years old are physically deteriorating, especially those which are consulted often and are therefore the most valuable.

Any member of the division including physicians, nurses and clerical staff has access to these records. There is no formal security for adding, updating and retrieving records. Subsequently, patient information, although philosophically confidential, is not protected by a formal security system.

The patient record is identified primarily by patient name, although the hospital patient identification number is recorded for billing and test-routing purposes. Problems with the name identifier include duplication, changing (through marriage and divorce for example), as well as mis-spelt names. For ease of access, the patient records are then ordered alphabetically by last name and arranged sequentially in the file.

Patient records are comprised from different data sources including:
1) letters of introduction from referring physicians,
2) previous records from other clinics, hospitals and laboratories,
3) the initial visit record, physical and clinic physician’s record,
4) succeeding visit and hospital records, medication and therapy changes,
5) laboratory data from tests ordered by the rheumatologists including tests sent to outside laboratories,
6) X-rays and records of all X-rays ordered by the rheumatologists,
7) indications of current clinic studies the patient may be participating in.

As a secondary case identifier, the current diagnosis is recorded on the outside of the patient file and subfiles. The files are also marked on the upper edge with diagnostic symbols; however, these symbols become damaged or removed with age. Continual handling of the record occurs on clinic days, when a patient is seen. At such time, the most recent records are removed physically and sent down to the clinic for the day. The records are then updated and returned to the files.

Patient records are not typically similar in format or structure creating a problem in data abstraction. Physician’s records whether from referring practitioners or University of Connecticut physicians reflect the personality of the physician. Even where, as in the clinic, the general layout is standardized, searching for specific information usually requires careful reading of the whole record.
Methods

System Design Considerations

The manually-accessed medical information system was reviewed with regard to both the types of data collected and the type of research and clinical studies conducted. With this information in mind, several different data forms were reviewed for applicability including: 1) prior data collection instruments used in research projects, 2) a general purpose database abstraction form from the South Carolina Medical Center, and 3) a data form used in a multi-center clinical trial involving patients with rheumatoid arthritis.

These forms were examined and evaluated by a group of rheumatologists for content, order and ease of use. Overall, they found the forms to be restrictive in content, unwieldy for comprehensive completion, difficult to understand and in atypical order. System analysts also evaluated these forms for 1) ease of data entry, 2) size of resulting data files and 3) general layout. The evaluators concluded that any data instrument used in a new data system must fit into the normal clinical pattern. Accordingly, the items should fit naturally into the usual clinic examinations. Also, in each of the above mentioned forms, there was a standard number of records for each patient and all records have to be completed for every patient at every relevant visit. However, the physicians require different types of information on different patients, and the number of studies completed on each patient varies from patient to patient. The database structure is therefore hierarchical in nature. For this system, proper case sorting and network variable links are needed.

Following this preliminary evaluation, examination of the operational flow of information in the clinic was undertaken, as well as survey of the general terminology and the research currently conducted in the Division of Rheumatology. It was essential to determine the type of questions the rheumatologists needed answered from their data for research purposes, and the type of information they need for clinical activities. By determining the needs of the division, relevant questions could be asked by database designers at forthcoming meetings.

System Designed Goals

Based upon the results of the preliminary evaluation, the preliminary goals of the new system were set as follows:

1) to provide a database form, in a logical sequence and clinically compatible for the rheumatologists. This form would become part of the medical record and replace many of the present bits of paper, especially the clinician's report. Because all patient information would be recorded in the same format irrespective of which physician completed the form, users searching for clinical information or conducting research would find their tasks much easier;

2) to acquire an interactive on-line computer system compatible with such a database with regard to size and structure;

3) to ensure that any on-line system was safeguarded from prying by unauthorized personnel,

4) to ensure that data in the database would be accessible to statistical packages used for analysis;

5) to provide a user-friendly means of data entry for non-computer personnel,

6) to provide a user-friendly means of browsing and data checking. This would allow rheumatologists to check entries to ensure their validity and would also allow browsing by clinicians. It would be preferable for clinicians interested in research topics to ask the computer to select for them records which contain select values or ranges of values for certain variables.

Technical Planning Stages: Data Collection Instrument

The development of the format for the data collection instrument involved a series of interactive meetings between a design team and users. The design team included a programmer-analyst, a data forms designer and a biostatistician. The users consisted of five faculty members in Rheumatology and several rheumatology fellows.

The data collection instrument was formulated according to the actual pattern of information flow. It contained several sections reflecting the accumulation of patient data from time of initial visit to subsequent follow-up visits.

At each main section; general variable groupings were constructed consisting of results of physical exams, history taking and test values. Subsequently at each meeting general outlines and contents were discussed. The most interested physician then offered to work further on the draft of variables and if necessary expand it. At later meetings, expanded copies of such drafts were circulated. The rheumatologists checked for missing items, discussed the order of items, and added further variable codes and variable labels.

After review and modification by the data forms designer, the data collection instrument was piloted by rheumatologists in the clinic. They brought the results of their test to the next meeting for discussion and further modification. This cycle continued until the data collection instrument was considered to be complete and comprehensive.

Proposed System: Computerized Rheumatology Information System

The final data collection instrument was comprised of several records which revealed the basic hierarchical design of the database system. The collection system consisted of: 1) an Initial Visit Record containing several sub-records such as physical exam and patient history, 2) follow-up visits consisting of clinic and hospital visits, and 3) special studies consisting of data from special research studies.

The Initial Visit Form contains over 1900 variables. It was designed to be completed by the attending physician and used directly by a data entry clerk in screen formatted mode. The main sections for the initial visit form are as follows:

1) Common information (for computer key),

2) Patient information (demographic),
Review of primary disease prior to initial visit to the clinic,
Constitutional symptoms (review),
Review of systems,
Family history,
Other past medical history,
Physical examination,
System examination,
Joint assessment,
Biopsies,
Laboratory Data Report Form (past tests and those ordered at initial visit).

Each main section was structured into major and minor variable headings. Minor headings were nested within major headings in such a fashion that if a major heading was negative, the physician would not be required to fill in minor headings. This design feature greatly facilitated rapid completion by the attending physician.

Follow-up forms are diagnosis specific for each main diagnostic category such as 1) rheumatoid arthritis, 2) systemic lupus erythematosus or 3) osteo-arthritis. The follow-up forms are much shorter in length than the Initial Visit Form and allow for a longitudinal approach of patient surveillance. In general, the follow-up forms reflect the more specific and sharper focus of disease-specific variables.

The Special Studies Forms are "study" or "investigator" specific. These forms are specially designed in a dynamic fashion each time a patient who is part of the major database is entered into a special research project.

Database Schema

The general database schema reflects the hierarchical structure of the database allowing for inclusion of varying numbers of follow-up visits and special studies among patients. The system also permits networking which allows follow-up records to be linked to diagnosis which is part of the Common Information Record (CIR).

Database Software

The software for the database management system, SIR (Scientific Information Retrieval), was selected because of its numerous utilities, including 1) portability to various computer systems, 2) its ability to assure the protection of the integrity of the data, 3) data storage and retrieval capabilities, 4) its hierarchical and networking capabilities, 5) form screen entry and 6) its ability to create SAVE FILES for common statistical packages (SPSS, BMDP and SAS).

Discussion

Consideration was given at several stages to protection of the system from tampering and prying. Computer privacy has been discussed in detail by Kroop and Prewitt. Here, our aim is to have the information easily available to those qualified to examine it, but to safeguard it from unauthorized tampering.

Password protection will be applied. The on-line database will be only available if a given password is entered into the computer. After passing this check, the user may browse or select specific records, but data entry or editing would require different passwords. Several different levels of access can be designated each with its own password. This is felt to be better than current checks on patient record use.

For record selection, sort identifiers are designated at each level. The Common Information Record will contain the main sort keys; patient identification number, diagnosis, date of diagnosis, patient status. Each of these variables can then be used for selection purposes. The date of each record or visit will always be one of the sorting variables. This will be both more reliable and faster than searching for information in the present system.

A complete description of sorting and protection capabilities are listed in the SIR manual and subsequent update manuals.

References


This paper is supported in part by Grant #AM20621, N.I.H., Multipurpose Arthritis Center.