COMPUTERS AND BIOMEDICAL RESEARCH 8, 279–293 (1975)

A Modular Self-Describing Clinical Databank System

STEPHEN WEYL

Artificial Intelligence Center, Stanford Research Institute, Menlo Park, California 94025

JAMES FRIES

Division of Immunology, Stanford School of Medicine, Stanford, California 94305

GIO WIEDERHOLD

Consultant, Stanford Medical Center, Stanford, California 94305

AND

FRANK GERMANO

University of Pennsylvania

Received April 29, 1974

Databanks containing detailed medical information have been established for specific divisions of the Stanford Medical Center using a table-driven computer system designed to record and analyze medical record data. The databank developed in the Division of Immunology over the past three years includes data from over 5500 patient visits. Nine months of experience in several other specialities has demonstrated the usefulness and generality of the approach.

The system offers administrative procedures to generate patient summaries and back up manual records, create graphical flowcharts, and print lists of selected patients. As an educational and research tool, the system provides "information services," including tabulations, graphical illustrations, and statistical analyses of medical variables. These services support patient care decisions by analysis of stored clinical experience. By increasing the information available to practicing physicians, medical speciality databanks may reinforce clinical practice sufficiently to allow modular expansion into a hospital network.

INTRODUCTION

In response to need for an inexpensive individualized clinical information retrieval system, a schema-driven database system (TOD) has been developed at the Stanford University Medical Center. The TOD (Time-Oriented Databank) system is the result of three man-years of database research and development undertaken collaboratively by physicians and computer scientists. A central concept for this system is the characterization of patients by means of numeric descriptors, recorded serially.

Traditionally, clinical description has been narrative and qualitative. When combining data from a group of patients, numeric representations are more efficient. Some clinical variables, such as "systolic blood pressure" and "blood urea nitrogen" are already quantified traditionally. Other variables, such as "shortness of breath" or "fatigue" may be described semiquantitatively using a standard five-point (0–4+) scale or other scales. Coding diagnoses with ICDA-8 nomenclature, and providing

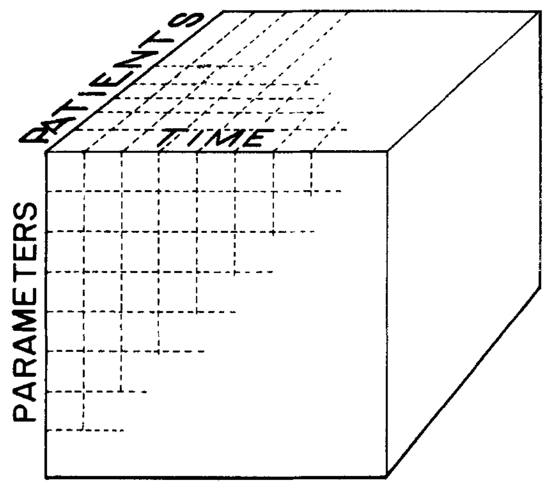


FIG. 1. The array of clinical data.

codes for interpretive material such as "chest x-ray" allow such data to be represented and managed numerically.

Numeric variables that characterize clinical observations are referred to as descriptors. The restriction that descriptors have numeric values results in loss of some information but is accompanied by major gains in reproducibility and standardization. Numerically coded observations lend themselves readily to computer analysis.

Three attributes are necessary and sufficient to uniquely describe a particular clinical observation: identification of the patient, the time, the name, and the units of the specific clinical descriptor. Some authors acknowledge attributes such as source, problem association, units of measurement, reliability, and others (*1*, *2*). However, by introducing conventions these auxiliary dimensions can be derived from the three primary attributes. For example, by convention "calcium" may be defined as a serological laboratory test measured in mgms percent by a particular technique.

Use of three attributes to define an observation allows conceptualization of clinical information as a three-dimensional array of data (Fig. 1). This conceptual framework can be applied in diverse medical specialties, and integrates readily with an interactive computer system. Such a system has been implemented with the following features.

(1) Prior to computerization of data in a specialty area a fixed format paper patient chart is designed by clinical experts in the area, ensuring practical, structured, and computer-compatible data recording. In most applications, a "time-oriented" flowsheet format has been chosen for recording of factual data. Problem-oriented progress note sheets urge the physician to make narrative descriptions of his clinical reasoning and plans; in most applications such information has not been computer stored.

(2) Each specialty group uses a separate databank. Data can be copied from one databank to another and information in one databank may be queried from another.

(3) Each databank is described in a standard language. The complete description of a databank is termed a "schema" (*3*, *4*) which provides control, explicit documentation, and linkage.

(4) A time-shared high-level language and file system make programming accessible to physicians. Several system features simplify the development of new procedures.

(5) An extensive library of public programs, available to all databanks, supports (a) data entry, verification, and correction; (b) creation of patient subsets based upon logical criteria; (c) administrative functions; and (d) information services including graphical presentation, tabulation, and statistical analysis of data.

(6) A master databank follows the course of all specialty databanks, maintaining a resource utilization profile for the collective.

These features are discussed below.

Manual Records as Retrieval Systems

As a first requirement, a medical information project must define rational concise procedures for the collection and review of data. The traditional disorganized medical record is not computer-compatible in a practical sense, and is not designed for efficient manual retrieval of data. A postulate of this project is that the paper patient record must be an information retrieval system in its own right, rather than solely an information receptacle.

The development of the databank begins with the design of a form for recording and retrieving pertinent information. This form is the primary medium that the medical staff interacts with while attending to the problems of the patient. To be both computer-compatible and clinically acceptable the form must enforce accuracy, thoroughness, conciseness, and logical organization. At the same time it must not impose a regimen so severe that creativity at the time of data-recording is diminished.

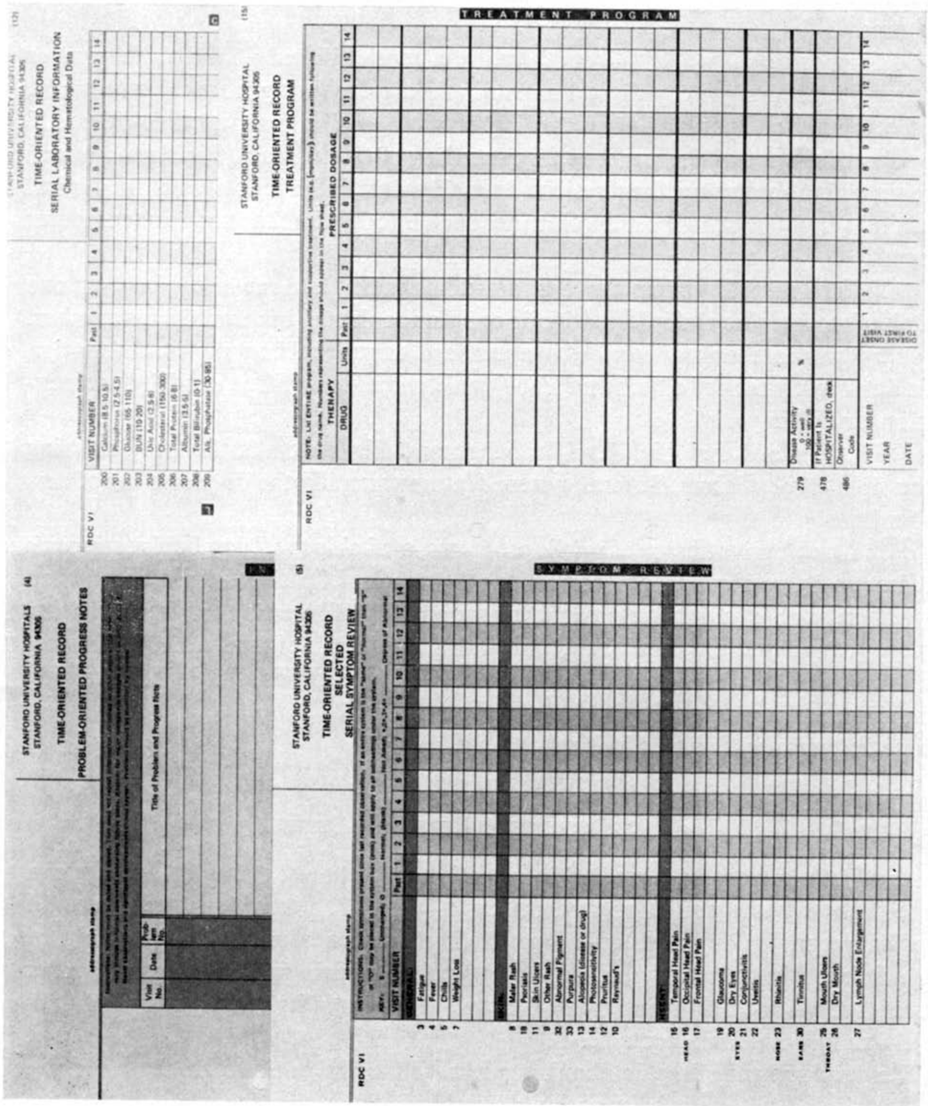
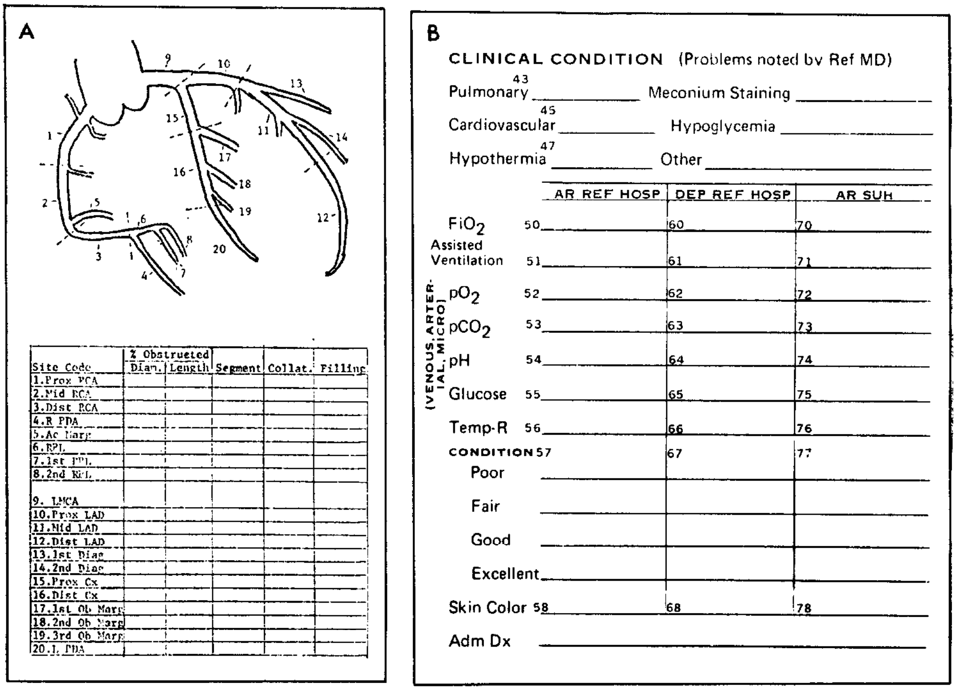


FIG. 2. Key pages of a time-oriented record (for the Division of Immunology).

Four years ago, we began to develop a medical record format subsequently termed the "time-oriented medical record" (*5*) for use in the Immunology Clinic at Stanford (Fig. 2). The high information density and logical organization of this chart, together with its acceptability to physicians led to the adoption of similar charts by clinics of Oncology, Metabolic Diseases, Cardiology, Endocrinology, Pediatric Orthopedic Surgery, Pediatric Rheumatology, and on a trial basis by a medical ward at the Palo Alto VA Hospital, as well as by other institutions and



Site Code	% Obstructed Diam.	% Obstructed Length	Segment	Collat.	Filling
1.Prox RCA					
2.Mid RCA					
3.Dist RCA					
4.R PDA					
5.Ac Marg					
6.RPL					
7.1st RPL					
8.2nd RPL					
9. LMCA					
10.Prox LAD					
11.Mid LAD					
12.Dist LAD					
13.1st Diag					
14.2nd Diag					
15.Prox Cx					
16.Dist Cx					
17.1st Ob Marg					
18.2nd Ob Marg					
19.3rd Ob Marg					
20.L PDA					

FIG. 3. Rather distinct fixed-format medical charts are appropriate for different specialties. Shown above are portions of forms used by Cardiology and Pediatrics for collecting computer data. (Permission of the Division of Cardiology and Pediatrics-Anesthesia, SUMC.)

individuals. Time studies have shown a marked increase in the efficiency of manual data retrieval using this format (*6*).

Different manual chart designs have evolved to meet the requirements of different databanks. Forms developed for Cardiology (Fig. 3A) record coronary arteriographic findings structured according to an anatomical diagram. Intensive Care Unit forms (Fig. 3B) collect data on infants treated by a Newborn Emergency Transport service. This form is designed to elicit rapid and complete responses over the telephone and in the transport vehicle.

Underlying each record design is the concept of clinical descriptors arranged in a fixed format. Observations are recorded as numeric values, numeric codes, or binary values. The full set of numbers recorded at a specific time represents the patient's status as an array of values, associated with a point in time. The use of a carefully designed fixed-format chart is satisfactory to physicians principally because of the improvement in manual information retrieval, exemplified by the ability to quickly visualize trends in a patient's course. Specific patient data may be rapidly reviewed, since it is stored in a standard location according to standard conventions.

After a medical specialty has defined effective patient recording and retrieval systems, entry of all or some portions of the data into a computer database system offers several advantages.

A file of carefully structured sorted paper documents is an excellent source for answering questions related to an individual patient. If the record is at hand, information may be extracted directly from a patient record in less time than is required to initiate a computer search request. However, a computer-readable file can afford major time savings when an inquiry requires a review of many patients. The computer analyses shown later in this report were generated in a few seconds at a unit cost of a few dollars.

Advantages of Specialty Organization

The quality of results obtained from a database depends on the quality of data contained in it. Quality control must be maintained at three levels: data included must be authoritative and clinically relevant; data recorded in charts must be consistent and complete; and transfer of data into the computer must be reliable.

Quality control at the data transfer point is relatively easily performed by the design of cautious interactive data entry procedures and error detection processing. The selection of appropriate clinical parameters and the enforcement of careful comprehensive recording is far more difficult. This type of quality control in a heterogeneous institution is best delegated to specific individuals and clinics with direct interest in the product. Definition of consistent and complete databases in specialty areas paves the way for national database definition (*7*).

Medical professionals are likely to be committed to projects over which they exercise authority. Placing databank responsibility in specialty areas allows individuals to be independently creative in the formulation of techniques, and encourages involvement in design throughout the medical community by the potential users themselves rather than confining system development to a small core of computer scientists and doctors.

Organization of a Modular Databank System

The computer may be given the specification of a fixed format manual chart. A computer-interpretable description of the chart is termed a "schema." A schema is initially defined in a high-level "databank description language" (*3*, *4*) which is translated into an internal form for machine processing. The schema specifies important attributes of each descriptor in a databank. These include a reference number, a description of contents and units of measurement, a short symbolic reference name, a data type, normal limits on value, and a default data entry value. The schema also specifies reformatting and data transformations to be performed by databank utility routines.

By using the schema file as primary documentation, and short symbolic reference names as links between identical descriptors in separate databanks, sharing of information between databanks becomes possible without major logistical overhead. A schema listing allows the sharing of standard descriptor lists between specialties, or between institutions.

Several provisions are made to stimulate hospital-wide development of the computer system. Retrieval modules, which are subroutines that obtain information from primary files, are provided to simplify programimng. Sample programs illustrate programming techniques to novices. A preprocessing program inserts internal procedure code into programs for standard functions such as patient subsetting, date conversions, and utilization accounting. Documentation at a number of levels is available to users.

An example of a retrieval module is the subroutine GPCOL (Get Parameter Column) which returns the array of descriptor information recorded on one occasion. Without a programmer being aware of internal file structure and linkage, this data record is retrieved efficiently with backward and forward pointers maintained to retrieve the next such record with minimal delay. Another example is the module GPEXT, (Get Parameter value in External Form) which returns a single descriptor element value in external form with dates converted and coded elements decoded. Again, sufficient linkage is maintained to minimize access time for further retrieval requests. In the context of the highly interactive PL/ACME system (*8*) writing databank programs is straightforward.

To provide rapid response when answering questions addressing data in different ways, auxiliary data files are maintained by utility programs. Primary data files are designed for optimal data entry, and are organized as a chain structure with successive visits of the same patient linked to each other (*4*). Auxiliary files include a "transposed" file which reformats data for access by element number.

New program features such as computation of dependent variables, special program options, new auxiliary files, or new retrieval modules are introduced in releases to the whole community. This is expedited by maintaining copies of all programs on a "TOD library." Programs in this library may be directly accessed by any PL/ACME user using a keyword after the standard "PROGRAM" statement. The library copy of a program is protected against modification by public users. Programs developed for specific groups which are generally useful are transferred to the public library.

HARDWARE AND SYSTEM SUPPORT

To develop a computer database system, an environment with adequate hardware and software is essential. At the Stanford University Medical Center, an IBM 370/158 computer with over one megabyte of primary storage and 300 megabytes of secondary (disc) storage is available. The configuration is equipped with a high-level time-sharing language PL/ACME which includes primitives for string manipulation and file access (*9*), simplifying development of physician-originated systems. Specific features which are desirable for support of a medical record database are dynamic allocation of file and record spaces, and the capability to extend records without copying files. Automatic compression of data spaces containing undefined, zero, or repeated data is important since observed values are sparsely distributed.

After a subspeciality schema has been defined, and an adequate amount of data has been entered into the computer using data entry modules, the system allows convenient access to the stored information.

THE PROGRAM LIBRARY: SECRETARIAL ASSISTANCE

Certain rather simple computer procedures are associated with the reformatting of data. These procedures are usually invoked by secretaries or other clerical personnel.

Figure 4 is a listing of certain attributes recorded for patients seen by a Newborn Emergency Transport service. After responding to a few simple prompts a listing is developed on a terminal describing babies admitted with below normal temperatures.

```
Enter subset number, or 0 to list index=?33
Subset 33 = TEMP < 35.5

     Column 1: Specify PARAMETER element=?age_xfer
     Column 2: Specify PARAMETER element=?birthwt
     Column 3: Specify PARAMETER element=?diag1
```

PATIENT'S COMPUTER RECORD NUMBER		AGE AT TRANSFER IN DAYS	BIRTH WEIGHT	PRIMARY DIAGNOSIS
TOR number	VISIT	AGE_XFER	BIRTHWT	DIAG1
417	1	.131250	454.545	1
413	1	.152083	3181.81	10
409	1	.229167	454.545	1
391	1	.194445	909.090	1
339	1	.339583	1534.09	1
300	1	.240277	1789.77	7
287	1	.208335	1306.82	7
278	1	.250000	985	7

FIG. 4. Age and birthweight characteristics of newborns admitted with below normal temperatures. User responses are underlined.

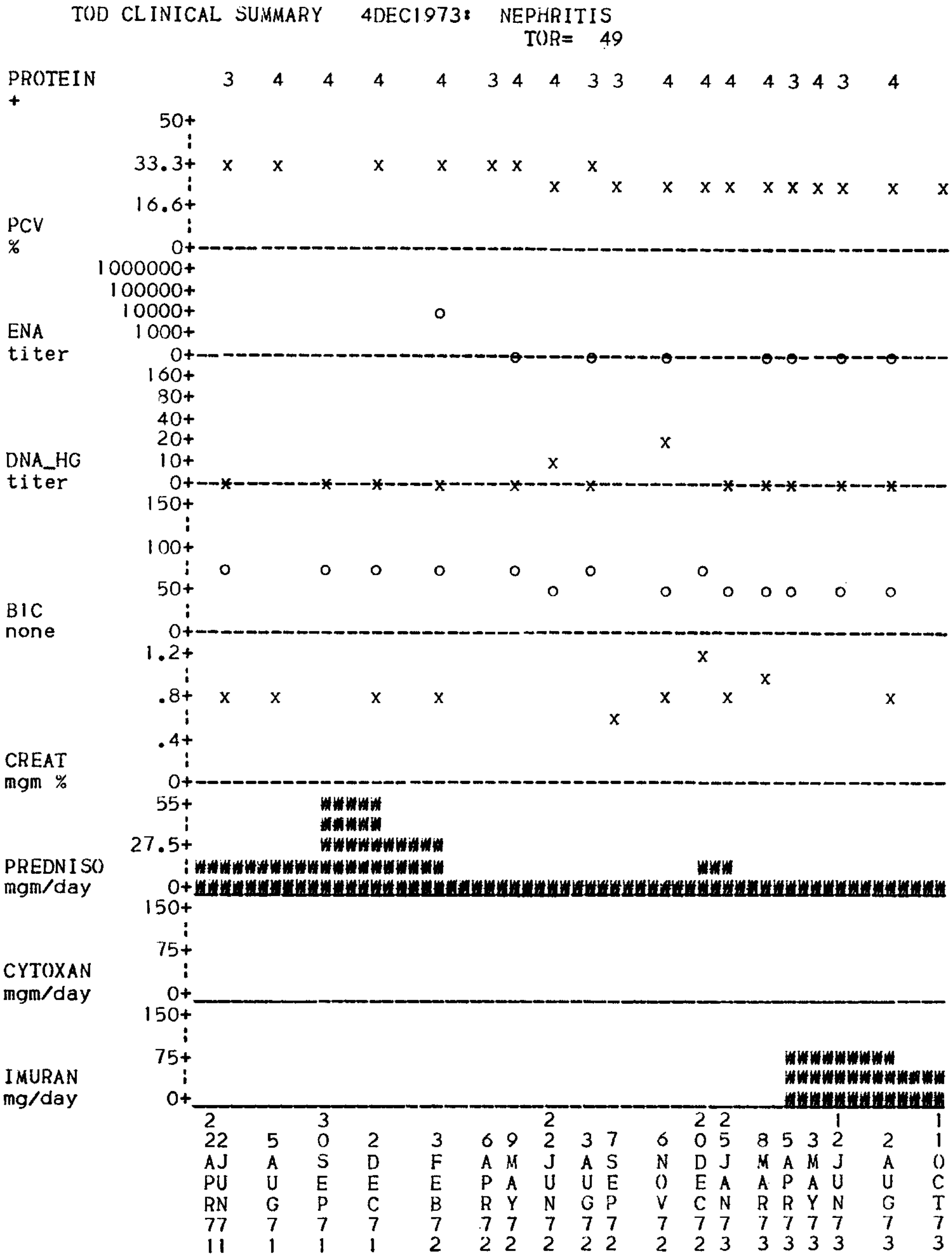


FIG 5. A computer-generated graphical flowchart.

The output in Fig. 5 is reformatting of information recorded in the patient chart. A flowchart characterizing certain aspects of a patient's course is selected from a group of possible formats and plotted by the computer. The resulting printout may be appended to the manual patient chart to provide the physician with a quickly absorbed multivariable review of the patient being treated. Figure 5 displays descriptors relating to nephritis for an Immunology patient.

The Program Library: Selection of Patient Subsets

A key problem when formulating expectations based on past experience is that of ensuring homogeneous patient populations for study (*10*). Classical diagnostic categories frequently contain very heterogeneous patient populations; this clinical variation limits the usefulness of the diagnostic category as a sole descriptor. The relation between the severity of skin rash and the magnitude of erythrocyte sedimentation rate observed in patients with lupus erythematosus breaks down if the patients have nephritis. A significant number of patients under 45 years of age undergoing immunosuppressant therapy may be expected to develop loss of fertility, a side effect not relevant to older patients.

Accurate clinical analysis requires the ability to construct homogeneous patient groups. Therefore, all tabulation and analysis programs begin by specifying a particular subset of patients for analysis. Patient subsetting programs perform this selection. These programs refine a master list of the current recorded patient population to a sublist. The sublist is then saved on a "library of subsets" for future use by analysis programs. Retention of sublists in a library allows a patient population to be analyzed by different programs or at different times without repeating the subsetting process. Subset lists are stored in the library with a catalog number, a user description, the date of creation, and a full description of the criteria used to select patients. Figure 6 illustrates the formation of a subset.

Subsets are defined by Boolean criteria. These criteria are associated either conjunctively or disjunctively. Boolean criteria are based upon demographic data (sex = MALE, birthday before 2AUG52) or descriptor data, (creatinine less than 1.5 mg%, creatine increasing by 0.5 mg%). The associated time dimension may be explicitly utilized when defining criteria. Valid descriptor criteria include: (1) descriptor value within specified bounds on a particular visit; (2) value increases (or decreases) to specified bounds; (3) value increases (or decreases) by a specified amount. The "and" procedure selects patients for which *all* criteria are satisfied. The user may specify that all criteria must be satisfied at the same point in time or that they may be satisfied at any point in the patient's course. The "or" procedure selects patients for which *any* criteria are satisfied, Subsets formed by using "and" and "or" procedures may then be complemented, intersected, or joined to create more complex subsets. Subsets may also be defined explicitly by entering a list of patient numbers. In this fashion, arbitrarily complex subsets may be obtained, while each step in the

```
FORM INITIAL SUBSET
Enter subset number, or 0 to list index=?2
Subset   2 = SLE,CLINICAL DX                                  ( 176 patients)
You will be creating Subset  67.
Number of criteria defining subset=?3
Test HEADER elements (Y or N)=?y
Include only those patients for which ALL criteria are satisfied (Y or N)=?y
Must criteria be satisfied simultaneously (Y or N)=?n
Use Visit 0 data (Y or N)=?y
Specify criterion  1:
     Element tested=?SEX
          low legal value=?0
          high legal value=?0
Specify criterion  2:
     Element tested=?PROTEIN
     Range 1, increase 2, decrease 3=?1
          low legal value=?1
          high legal value=?4
Specify criterion  3:
     Element tested=?SED_RATE
     Range 1, increase 2, decrease 3=?2
     Increase to range (Y or N)=?n
     Amount of increase=?5
     1 for change across all visits, 0 for change between two visits=?0
Title of subset: =?FEMALE SLE PATIENTS WITH  PROTEINURIA AND INCREASING ESR

Subset  67 = FEMALE SLE PATIENTS WITH  PROTEINURIA AN  (  36 patients)
 Criteria selected from transposed files.
  Formed from Subset   2.  Defined by the following  3 event(s):
  Patients with all criteria satisfied at some time.
  Visit 0 data included.
    Header element 11:  range from 0 to 0
    Parameter element 269:  range from 1 to 4
    Parameter element 236:  increased by 5
```

FIG. 6. Forming a subset on the computer. User responses are underlined. At the immunology clinic, what SLE patients had proteinuria and increasing ESR?

process remains elementary and expressible by an interactive dialogue meaningful to the untrained user.

THE PROGRAM LIBRARY: UTILIZATION OF STORED EXPERIENCE

Clinical decisions are based upon expectations. Well informed expectations may derive from a large and accurate base of past experience. Tabulation of experience allows derivation of realistic expectations. The program library offers a variety of procedures for analyzing experience. These procedures provide information for the clinical decision process.

One characteristic of a descriptor is the distribution of its values. The distribution may be normal or it may take other forms. Figure 7 summarizes the distribution of age for systemic lupus erythematosus from the Immunology databank. A skew in the tail of the distribution is seen. In addition, this program provides a statistical analysis of the shape of the curve.

Another attribute of a descriptor is its mean. A typical question is: "What values of common hematology tests are usual in SLE?" When values of a descriptor are normally distributed, the standard deviation is another useful measure (Fig. 8).

Specific questions about the magnitude of abnormalities noted on a group of patients may then be asked, and deductive inferences made.

Other procedures allow an estimate of the prognosis for a subpopulation of

```
CAN A 60 YEAR OLD PATIENT HAVE SLE?

Subset  2 = LUPUS                                        ( 165 patients included)
FREQUENCY DISTRIBUTION FIRST VALUE PROFILE FOR AGE       (   2), years
Interval/Value   Pts.     %  !
      < 0.0        0     0.0 !
   0.0- 6.7        0     0.0 !
   6.7-13.3        7     4.3 !****
  13.3-20.0       21    12.8 !*************
  20.0-26.7       33    20.1 !********************
  26.7-33.3       28    17.1 !*****************
  33.3-40.0       22    13.4 !*************
  40.0-46.7       16     9.8 !**********
  46.7-53.3       19    11.6 !************
  53.3-60.0       10     6.1 !******
  60.0-66.7        5     3.0 !***
  66.7-73.3        2     1.2 !*
  73.3-80.0        1     0.6 !*
  80.0-86.7        0     0.0 !
  86.7-93.3        0     0.0 !
  93.3-     100    0     0.0 !
      >     100    0     0.0 !
TOTAL            164   100.0 I.........I.........I.........I.........I.........I
                             0        10        20        30        40        50
   ! undefined points! Mean =     34.0! Standard Deviation =      14.60661
```

FIG. 7. Age distribution for LUPUS Patients, produced by a histogram program.

```
WHAT VALUES OF COMMON HEMATOLOGY TESTS ARE USUAL IN SLE?

 TOD SYMPTOM FREQUENCY REVIEW                All values
Element    %pos    # found    Mean    Std.Dev    # > zero    Mean      Std.Dev
 WBC       100.0    1110      6.699    3.264      1110
 PCV       100.0    1097     36.333    6.929      1097
 HEMOGLOB  100.0    1032     12.026    2.324      1032
 NEUT      100.0    1017     70.826   15.020      1016      70.896    14.862
 BANDS      63.0     899      2.933    4.691       563       4.684     5.190
 LYMPHS    100.0    1016     19.158   12.364      1013      19.215    12.338
 MONOCYTE   95.0     999      4.901    3.661       950       5.154     3.577
 EOSIN      58.0     878      1.378    1.932       511       2.368     2.018
 BASOS      30.0     791      0.436    0.914       235       1.468     1.138
 PLATELET   99.0     306    260.577  137.092       304     262.292   135.898
 RETICUL   100.0     197      2.221    2.416       197
 SED_RATE  100.0     901     36.345   18.106       901
   9 patients died,      5 patients were lost to followup.
```

FIG. 8. What values of common hematology tests are usual in SLE?

patients with a particular syndrome, or undergoing a particular therapy. The summary shown in Fig. 9 incorporates a generalized concept of prognosis discussed in detail in another paper (*11*). Prognostic summaries may be requested by physicians encountering a particular clinical situation.

These tabulations were generated in less than two minutes of terminal time after only a few seconds of computer search and computation. The usefulness of such analyses based on years of recorded clinical data is particularly evident when contrasted with the quality, magnitude, and availability of similar information from the medical literature.

```
PROGNOSTIC ANALYSIS     SLE,CLINICAL DX

Starting event: 70 patients.                 Endpoint defined for 19 patients.
PLEURISY between 1 and 4 (+)                 CREAT increased by .200000 (mgm %)
```

TIME MONTHS	ENDPOINT	LOST	FOLLOWED	PROB. ENDPOINT	% NOT REACHING ENDPOINT	STANDARD ERROR
1	3	8	70	0.05	95.45	2.56
2	3	2	59	0.05	90.52	3.69
3	2	0	54	0.04	87.16	4.25
4	2	1	52	0.04	83.78	4.71
5	1	0	49	0.02	82.07	4.91
6	1	1	48	0.02	80.34	5.10
7	0	4	46	0.0	80.34	5.10
8	0	2	42	0.0	80.34	5.10
9	1	0	40	0.02	78.33	5.36
10	0	0	39	0.0	78.33	5.36
11	0	1	39	0.0	78.33	5.36
12	0	1	38	0.0	78.33	5.36
13	1	2	37	0.03	76.16	5.63
14	1	1	34	0.03	73.88	5.91
15	2	1	32	0.06	69.19	6.40

9 patients died, 5 were lost to followup.

FIG. 9. Following pleurisy, what is the probability of renal deterioration in SLE? This "life table" analysis shows the decreasing percentage of patients maintaining stable renal function (serum creatinine not increasing by 0.5 mg %) over a period of months.

THE PROGRAM LIBRARY: THE COMPUTER AS ILLUSTRATOR

The impact of visual data is largely due to its immediacy and density. The flowchart shown in Fig. 5 rapidly conveys the interrelationships between antibodies to DNA and subsequent change in renal function. Extraction of this overview from an array of numerical data would require more effort.

A scatterplot immediately displays many data points. This provides an investigator with more than a summary of data; he obtains a picture of *all* data in a manner which allows visual correlation and interpretation. A computer generated scatterplot of the relationship between aspirin dosage and serum uric acid values is shown in Fig. 10. In this scatterplot a time delay of one visit is introduced between the prescription of aspirin and its observed effect. Other program options allow selection for each patient of the first or last recorded x, y pair, all recorded pairs, the y corresponding

```
249 data points tabulated
  1 data points encountered which were out of bounds
A     30.00I          2
S          I
A          I          1
           I          1
=          I
      20.00I       11211 2 111              1
t          I       23111
a          I       12  1  2 31       1 1 11         1
b          I       2 22112113                 1
s          I        1135441531 113  2                          1
/     10.00I          1     11   1 2   1  1   1
d          I          2 11   1 1   11         1
a          I            1 21111 1   1   1  1 1      1
y          I        1 11     11 1       1  1
           I       1 21   11 1  21 13 11 11 11
      0.0  --------131253418254744137343 44-2243-1----11---1-----------
           I          I        I         I        I         I         I
           I          2.00     I         6.00     I         10.00     I
          0.0                  4.00               8.00                12.00
RHEUMATOID ARTHRITIS                                  UR_ACID , mgm %
LINEAR REGRESSION ( 249 data points collected):
Estimated y=     14.1917 +     -1.8907 X.  Correlation= -.4645.
```

FIG. 10. Can aspirin therapy result in significant hypeuricemia?

to the highest or the lowest value of x, and the x corresponding to the highest or the lowest value of y. Values for x and y within a pair are always recorded simultaneously.

A Databank to Describe Databanks

The consistency of databank implementation among the several clinics using the TOD system leads naturally to the formulation of descriptors that monitor the progress of individual databanks. Variables measure the cost of utilizing different modules, the rate of growth in file storage, the density of recorded data, the data entry error rate, and the utilization of auxiliary files. A DATABANK databank has been developed to follow these variables, supporting predictions of cost and resource utilization for specific projects.

To facilitate utilization analysis, all databank programs record certain information during their runs. This information is summarized in a cost report sent to databank supervisors each month, providing a profile of costs and utilization. Certain information from these reports is entered immediately into the DATABANK databank. This database may be used to substantiate decisions related to system costs and the optimization of specific modules.

Discussion

The TOD system is a fully operational set of programs that has been available since April 1973 on the Stanford University Medical Center facility. The full system consists of more than 60 programs which support data entry and update, file definition and maintenance, and analysis functions. Full documentation for the system in the form of a user's guide (TOD Manual) may be obtained from the Stanford Center for Information Processing: Medical Computing Services.

Key features of the TOD system which have encouraged its acceptance by diverse clinics are its fixed-format manual charts, its decentralized approach to database definition, and its provision of a mode for collaborative design of databank procedures.

The descriptor concept central to TOD simplifies the computer system, without placing major restrictions on the nature of data that may be processed. Textual information may be included and is available for narrative output procedures, but not for tabulation.

Large quantities of local, standard data readily available for analysis may have a marked effect upon medical research and education. Large scale databank facilities are likely to impact the practice of medicine significantly. Successful, logically interrelated, syntactically compatible databank modules offer promise for linkage into viable institutional systems.

REFERENCES

1. BRUNJES, S. An anamnestic matrix toward a medical language. *Comput. Biomed. Res.* **4,** 571–584 (1971).
2. HURST, J. AND WALKER, K. "The Problem-Oriented System." MEDCOM Press, 1972.
3. "Report to the Data Base Task Group." CODASYL Data Base Task Group, ACM, April 1971.
4. WIEDERHOLD, G. Database design and schemas. Class notes for MIS-290. Department of Medical Information Science, University College of California at San Francisco.
5. FRIES, J. F. Time-oriented patient records and a computer databank. *JAMA* **222,** 1536–1542 (1972).
6. FRIES, J. F. Alternatives in medical records. *Medical Care* **12,** 871–881 (1974).
7. FRIES, J. F., HESS, E., AND KLINENBERG, J. A. A standard database for rheumatic disease. *Arth. Rheum.*, **17,** 327–336 (1974).
8. WIEDERHOLD, G. An advanced computer system for medical research. *Proc. of the IBM Japan Computer Science Symposium-Research and Development and Computer Systems.* Tokyo, Japan, pp. B1–B15 (1969).
9. FREY, R., GIRARDI, S., AND WIEDERHOLD, G. A filing system for medical research. *Bio-Medical Computing* **2,** (1971).
10. FEINSTEIN, A. R. "Clinical Judgment." Williams & Wilkins Co., 1962.
11. FRIES, J. F., WEYL, S., AND HOLMAN, H. Estimating prognosis in systemic lupus erythematosus. *Amer. J. Med.* **57,** 561–565 (1974).