Man-machine communications in the biological-medical research environment

by WILLIAM E. FARLEY, ALFRED H. PULIDO, TATE M. MINCKLER, and LEE D. CADY, JR.
The University of Texas M. D. Anderson Hospital and Tumor Institute
Houston, Texas

The key source of raw data in most biomedical research is the patient's medical record. The hospital patient medical record is most commonly thought of as the repository for all pertinent facts relating to laboratory test results, diagnostic conclusions, treatment procedures, and observations. Depending on the nature of the patient's complaint, there are varying amounts of medical history information incorporated into the record. In many instances, the compilation of the record has become so routine that little thought is given to its potential value as a medical research tool.

Ideally, the medical record would serve the clinician during treatment of the patient and when no longer needed by him would pass to the researcher to be associated with other similar records for statistical analysis. Obviously, the method presently used would satisfy both of these needs to only a limited extent.

In recent years, many methods have been designed to record and store patient information in a more usable fashion for both the clinician and the researcher. A number of these methods have proved successful.[1,2] The extent to which any of these systems have succeeded appears to be directly related to two factors: (A) simplicity of recording and retrieving the information, and (B) system operation cost[3] in terms of dollars and physicians' time.

Maximum utility is derived from those systems which incorporate both retrospective and prospective data accumulation. Unfortunately, most of the information accumulated in old records is inaccessible to the researcher because of the lack of available trained personnel to extract the information and the associated expense of manual conversion.

Since the institution was founded in 1947, well over 60.5 thousand medical records have been compiled on cancer patients treated at The University of Texas M. D. Anderson Hospital and Tumor Institute in Houston, Texas. Specific information used in individual research projects is manually extracted at various times, but this constitutes only a very limited subset of the totality of available recorded information.

The possible rewards of a totally automated medical records system at M. D. Anderson Hospital were carefully weighed against the negative considerations of cost and time spent which potentially would yield little of value. As a result of this comparison, the decision was made to initiate a pilot project directed toward automating a portion of the medical record. The outcome of the pilot study would then form a basis for determining the practicality of continuing the project.

Traditionally, the examining physician follows a fairly well-defined procedure in performing the initial examination of a new patient. In most instances, a standard form is utilized. This form is used not only to record needed information, but also as a checklist for the physician to insure that all relevant data is obtained. Upon completion of the examination, the physician usually amplifies his findings with narrative comments. Laboratory tests may be ordered to aid in reaching a diagnostic conclusion or to substantiate a tentative diagnosis. Additional tests such as electroencephalograms, electrocardiograms, urinalyses, chemical, bacteriological and hematological analyses must also become part of the total patient record, contributing to the complex of data which must be dealt with.

Our initial efforts were directed toward the original data element or the preliminary examination form with its associated comments. Implementation was on an IBM 1401 16K system configuration (Figure 1) which was well suited to our purpose. Having a console inquiry station, we designed our program for optimum man-machine interaction.

"One of the most important problems in the area of computer storage and retrieval of clinical information concerns the types of formats used for obtaining and selecting information."[4] Code sheets designed specifically for transcription of previously recorded data were used. Each line of the sheet represents one 80-column

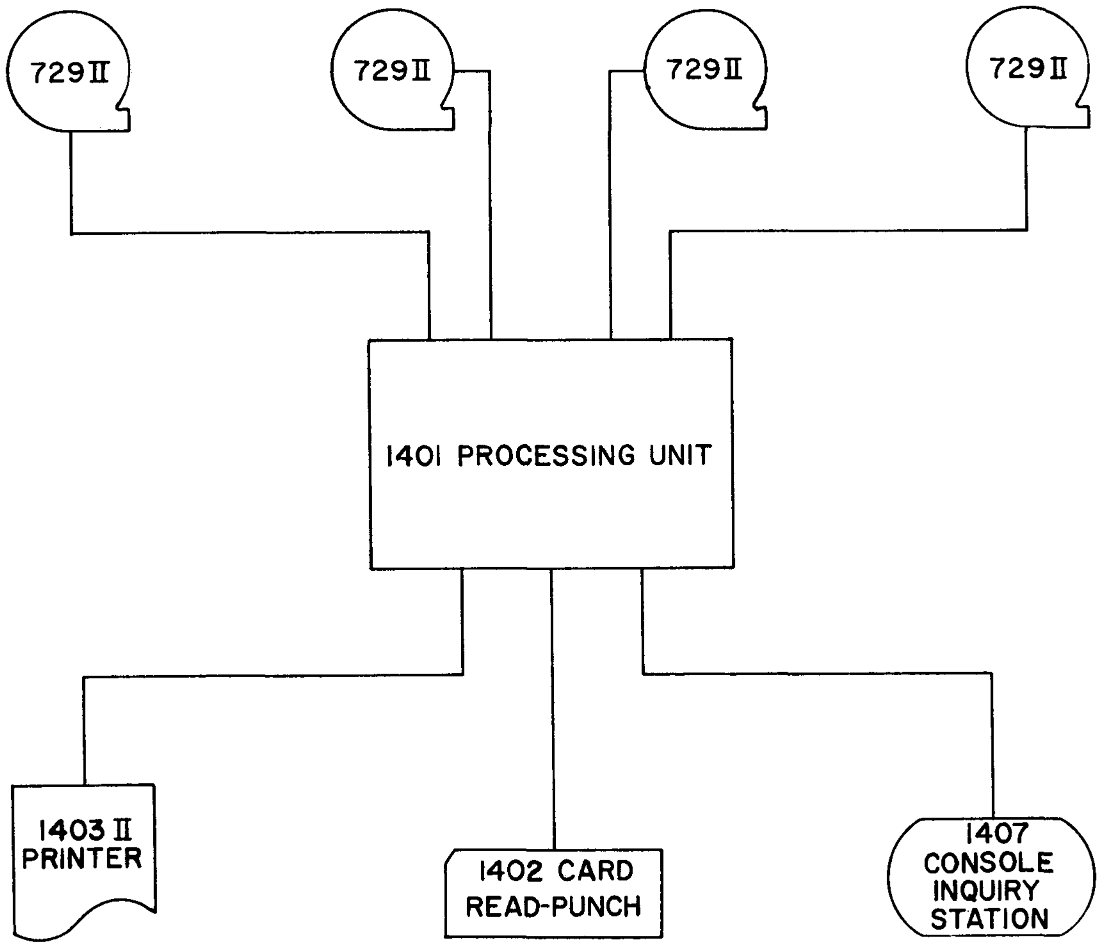

Figure 1 — IBM 1401 16K system configuration

Field Name	TYPE CODE	PATIENT IDENTIFICATION	DATE	CARD NO.	CONTINUATION
Size	X	XXXXXXXXX	XXXXXX	XX	XX
Card Column	1	10	16	18	20

Figure 2 — Fixed-field format identifiers

```
247AAJ  PM PHYSICAL MEASUREMENTS
247ABJ    HEIGHT  WEIGHT  HAIR COLOR  EYES COLOR  BUILD    TEMPERATURE
248AAJ    BLOOD PRESSURES
248ABJ     SITTING    RECUMBENT    STANDING
248AC/    SYS. DIAS.  SYS. DIAS.  SYS. DIAS.
249AAJ    PULSE
249AB/              AFTER     2 MIN.                AFTER STANDING
249AC/     SITTING  EXERCISE  AFTER    RECUMBENT    3 MINUTES
```

Figure 3 — Heading records

punch card. Fixed-field format identifiers are punched into the first twenty columns of each card (Figure 2).

As the preliminary physical examination of most patients contains the patient's medical history in addition to his examination results, these elements were chosen as our first two data divisions. Recognizing the potentiality of multiple major divisions of this type in the total record, column one is allocated for a numeric

Field Name	TYPE CODE	CARD NO.	CONTINUATION	CARRIAGE CTL.
Size	X	XX	XX	X
Card Column	1	3	5	6

Figure 4 — Control information

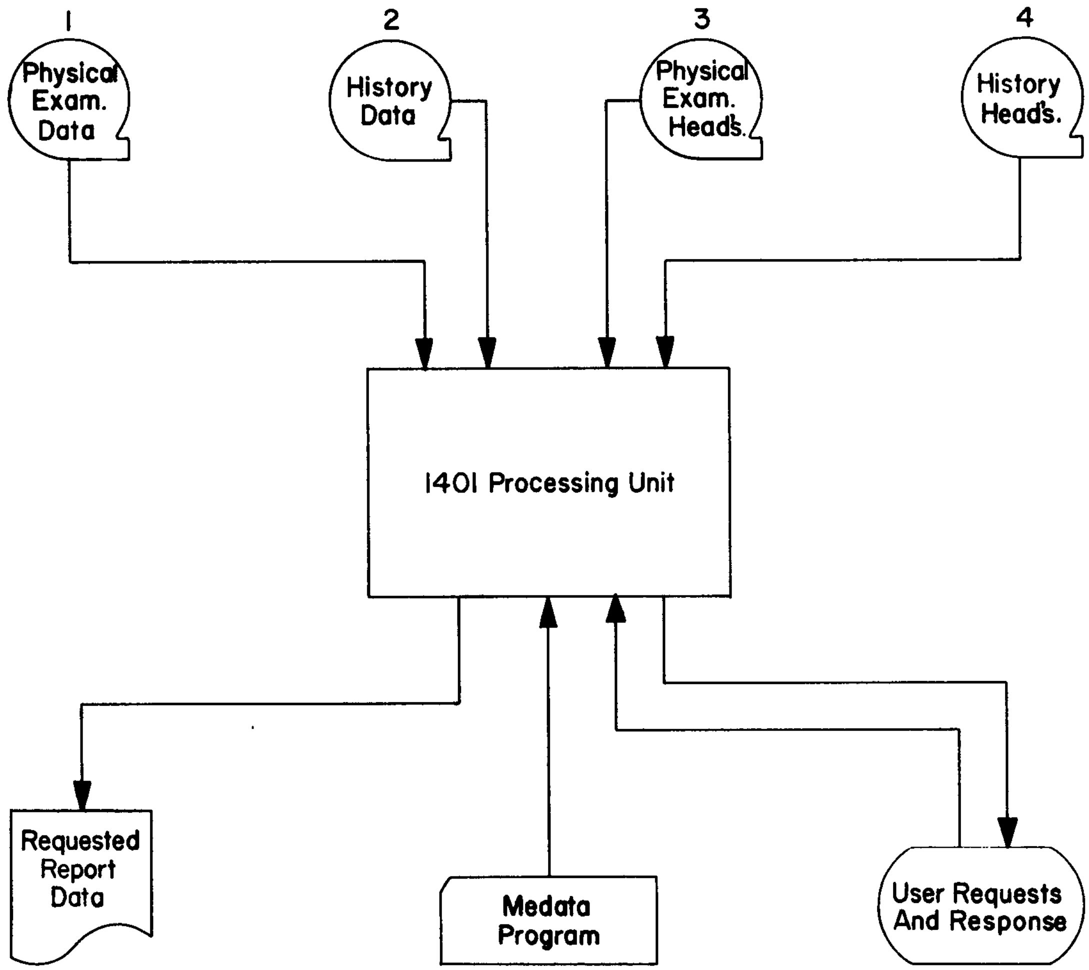

Figure 5 — Program unit assignments

code identifying each division. Adequate space is provided for patient identification number and record date. Card number serves a twofold purpose in that it is a multipurpose sequence control and format designator for the remaining sixty columns of each card.

Heading records (Figure 3), with card numbers corresponding to each data record format and ordered in ascending sequence, are available to the program for report generation. The first six characters of each heading record are reserved for control information (Figure 4).

Operationally, the program relies on user input of search parameters from the console inquiry station. Figure 5 shows the unit assignments for the program. At present, examination information (Tape Unit 1) and history information (Tape Unit 2) are on separate tapes. However, expansion of the system would necessitate combining these and other data into multipurpose tapes. This would also be true in the case of heading records (Tape Units 3 and 4).

When loaded, the program accesses the console inquiry station causing the question "WHO" to be

```
WHO
000201161 (USER RESPONSE)
WHAT
PM        (USER RESPONSE)
WHEN
15 FEB 65 (USER RESPONSE)


    ** RECAP **

WHO 000201161 BANKSTON, FRED

WHAT PM PHYSICAL MEASUREMENTS

WHEN 15 FEB 65


IF CORRECT PRESS START IF NOT PRESS REQUEST ENTER KEY THEN PRESS START
```

Figure 6 — Console inquiry log

CODE	DESCRIPTION	CODE	DESCRIPTION
CC	PRESENT COMPLAINT/ILLNESS	LB	LABORATORY FINDINGS
ID	IDENTIFICATION INFORMATION	MD	EXAMINING PHYSICIANS
DN	DENTAL	NP	NEUROPSYCHIATRIC EXAMINATION
DX	DIAGNOSIS	PE	PHYSICAL EXAMINATION
ER	EAR -HEARING-	PH	PERSONAL HISTORY
EY	EYE	PM	PHYSICAL MEASUREMENTS
FH	FAMILY HISTORY	SR	SYSTEM REVIEW
HX	MEDICAL HISTORY		

Figure 7 — Group item response codes

typed, then halts. To initiate a search, the user responds with a patient identification number (Figure 6) which is validated by the program. Should a valid number be entered by the user, the second question "WHAT" is asked by the program. User response may be any one of the two-digit group item response codes (Figure 7). The item designation will be checked by the program for validity before the final question "WHEN" is asked. User response to this question may be either a single date or a date range comprised of initial date and last date. The format used is the standard military date designation. Consistency checks are made by the program on the date or dates entered before a recapitulation of the total request is made on the Console Inquiry Log (Figure 6). The search may then be initiated by pressing the start key. When found, the requested information in report format (Figure 8) is listed on the line printer.

In the final analysis of the pilot project, we proved to ourselves that the system is practical, notwithstanding the expense of manually coding the information to be processed (average time required to transcribe one

```
IDENTIFICATION NUMBER 000201161   NAME BANKSTON, FRED

 PM PHYSICAL MEASUREMENTS

   HEIGHT   WEIGHT   HAIR COLOR   EYES COLOR   BUILD     TEMPERATURE
   70.00    172.00   BLOND        BLUE         HEAVY      98.0

   BLOOD PRESSURES

    SITTING     RECUMBENT     STANDING
   SYS. DIAS.   SYS. DIAS.   SYS. DIAS.
   130   84     128   88     130   88

   PULSE
              AFTER      2 MIN.                 AFTER STANDING
    SITTING   EXERCISE   AFTER     RECUMBENT    3 MINUTES
     68       100         72       68            80
```

Figure 8 — Requested information in report format

physical examination is approximately one hour) and the associated system operating cost. Core storage requirements are high (12K +), and output mediums are not flexible enough to satisfy future demands which we anticipate.

The experience gained from this pilot study is a valuable tool for use in projecting the future hardware environment in which any successful large-scale system of this type must exist. One may readily grasp the impracticality of attempting to store 60.5 thousand medical records, with an average thickness of over three inches, on magnetic tape with any expectation of speedily searching them serially. Therefore, random access mass storage capability having a reasonable transfer rate must be provided. Under no circumstances could economic justification be found to support the idea of a one user main-frame of even the smallest design; consequently, a main-frame designed to accommodate multiple simultaneous users with their corresponding I/O terminals must exist. The configuration of these terminals may vary depending on the user needs.

In the absence of this "ideal" hardware environment, our immediate plans are to expand the present system to include digital editing and file maintenance of recorded analog signals from various laboratory procedures (EKG, EEG, etc.).[5] To accomplish this, we are implementing this basic system on an SDS 930 Computer, which was recently acquired. To speed up data recording, an IBM 1050 system will be used.

Our long range goals are to make available to the research physician the wealth of information now buried in the medical records of patients previously hospitalized as well as to furnish the clinician immediate access to all pertinent information on a patient under his care.

REFERENCES

1 J E SCHENTHAL J W SWEENEY W J NETTLETON JR R D YODER
Clinical application of electronic data processing apparatus — III
System for Processing of Medical Records JAMA 186 pp 101-105 1963

2 R P LEVY M R CAMMARN M J SMITH
Computing handling of ambulatory clinic records
JAMA 190 12 pp 77-81 1964

3 H W BAIRD J M GARFUNKEL
Electronic data processing of medical records
N Eng J of Med 272 23 pp 1211-1215 1965

4 J KOREIN L J TICK M A WOODBURY L D CADY JR A L GOODGOLD C T RANDT
Computer processing of medical data by variable-field-length format
JAMA 186 pp 132-138 1963

5 L D CADY JR
Screening cardiograms by electronic means
Trans Assoc Life Ins Med Dir of Amer 48 pp 96-117 1964