COMPUTER MANAGEMENT OF CLINICAL INFORMATION: CAPTURE AND RETRIEVAL OF CLINICAL ORTHOPEDIC DATA BY MEANS OF THE VARIABLE-FIELD-LENGTH FORMAT*

RALPH LUSSKIN, M.D., JULIUS KOREIN, M.D., WALTER A. L. THOMPSON, M.D., AND THOMAS HEFFERNAN

Departments of Orthopedic Surgery and Neurology
New York University School of Medicine
New York, N. Y.

THIS report describes the systematic capture, storage, and retrieval of orthopedic clinical information by means of keyboard devices and digital computers at New York University Medical Center and Bellevue Hospital Center, New York. Of prime interest to both academic and clinical surgeons is the fact that this information is generated as a by-product of routine medical practice and that it does not involve coding of diagnoses, procedures, or other clinical information. Once these narrative reports are prepared for processing by the computer their content can be analyzed almost without restrictions. Thus all routine data collection, i.e., the presently contracted inpatient or outpatient clinical record, proceeds as prior to the application of this system.

Any orthopedic clinical facility may now begin to order the routine preparation of information for later insertion into the computer. In addition, all such facilities could generate compatible documents which could be processed in a regional or national center without the introduction of laborious or restrictive coding.

Possibilities include not only the collection of statistics of disease, the evaluation of treatment, and the control of complications, but extend

*Presented at a combined meeting of the Section on Biomedical Engineering and the Section on Orthopedic Surgery of the New York Academy of Medicine held at the Academy on February 17, 1972.

This study was supported in part by Bellevue Hospital Medical Center, New York, N.Y., by Grant No. I-470 from the Health Research Council of the City of New York; Grant No. PH 43-67-1336 from the National Cancer Institute, Bethesda, Md.; Grant No. 8-0164-708 from the Neurology Research Fund, New York University Medical Center, New York, N.Y.; Grant No. 605 from the Maternal and Child Health Service, Integrated Comprehensive Children and Youth Project, Department of Health, Education, and Welfare, Washington, D.C.; and Public Health Service Grant No. 00126 from the National Center for Health Services Research and Development, Washington, D.C.

to simplified procedures for utilization review and audit. These latter have the potential for imposing great burdens on the medical community. The possibility of simplifying these tasks by using computer techniques to "study" the medical record should be of great interest.

All these tasks require the ordering and special handling of medical information when it has been recorded. This paper will describe the capture of clinical information, its insertion into a computer facility, and its subsequent retrieval for analysis and for the care of patients.

Introduction

The modern digital computer can process numerical data at high speed. It can store information, modify and augment stored information, and compare new information with old. It can follow directions and it can modify its own operating instructions under special circumstances. Numerical, business, and production tasks have been processed efficiently by computers for 20 years and there has been great impetus to apply this equipment to medical tasks.

Since the capacity of the newer computing devices is vast and their speed incredible (millions of operations per second), nonnumerical information can be processed also. As each letter of the English or any other natural language can be given as a numerical symbol, entire words or even combinations of words may be represented by numerical sequences. Each of these numerical sequences represents a specific word or words and can be identified when compared with an internally maintained word list.

The physician must order his words so that information can later be searched out in a logical way. While the operation of computers is in itself of great interest, the clinician must not become overly entranced with these devices. He must carefully analyze and emphasize the goals to be achieved and their requirements in terms of time, effort, and money, lest the "rapture of the deep" engulf him as it has so many others. Once the computer operation itself becomes the dominant interest of the physician, his costs will rise as his success recedes. We shall allude to the internal operations of the machines used in this work, but one must avoid excessive identification with these devices.

There has been rapid change in technology and no doubt this will continue. Computers, peripheral devices ("hardware"), programs, and techniques of capture will all change to meet the needs of a rapidly

FIGURE I. FORMAT FOR ORTHOPEDIC DISCHARGE SUMMARY

¼

(ODS)

IDENTIFICATION NO. —

BELLEVUE HOSPITAL — ORTHOPEDIC SURGERY — DISCHARGE SUMMARY

NAME — SEX —

DATE OF ADMISSION — DATE OF DISCHARGE —

LOCATION —

BIRTH DATE —

AGE — RACE —

SURGERY —

PM —

##

(HISTORY)
CHIEF COMPLAINT —
PRESENT ILLNESS —
PAST HISTORY —
FAMILY HISTORY —

(PHYSICAL)
VITAL SIGNS —
GENERAL PHYSICAL
ORTHOPEDIC EXAMINATION —
NEUROLOGICAL EXAMINATION —

(LABORATORY FINDINGS)
URINE —
BLOOD —
X-RAYS —
OTHER LAB —

(COURSE)
COURSE IN HOSPITAL —
COMPLICATIONS CODE —
STATUS AT DISCHARGE —
DISPOSITION —

(DIAGNOSIS)
ADMISSION DIAGNOSIS —
FINAL DIAGNOSIS —
SECONDARY DIAGNOSIS —

(KEY DATA)

(DICTATED) —

++++

changing environment. The clinician must direct his interest to his own work and to the actual results which can be delivered. To act otherwise is to court failure, as many highly respected businessmen have learned.

Much effort has been expended in using the computer for diagnostic purposes. In orthopedics, as in most other medical fields, the reliability or weighted value of each piece of evidence is so subjective that numerical evaluation appears to have limited potential. We state this in order to contrast diagnostic computer procedures with the relatively prosaic task of record keeping. It is the capacity to manage a tremendous volume of medical records and to store these records that necessitates and justifies the introduction of the digital computer.

Consideration should now be given to the various general solutions to the problem of information processing and analysis which have been attempted previously.

The first and perhaps the most widely practiced method of storing information is coding. This is derived directly from the practices of the last century. At that time the ordering of diagnoses into specific categories permitted the collection of a great quantity of clinical data. As our experience, skill, and knowledge increased, it became increasingly difficult to fit the available information into categories. The labor involved in the attempt has become a major impediment to the accuracy of the process. The unknown or nonclassifiable portion of the data begins to outweigh the classifiable material. When circumstances or approaches change, the coded data are insufficient to satisfy the inquiry and a laborious return to the original and often more undecipherable clinical record is required.

Check list restricted-data formats are described by Seed[14] and Brolin.[1] All such procedures severely limit the amount of information which can be retained in the system and impose odious constraint on the physician. The same can be said of data-capture methods which utilize a physician-created dictionary and hierarchy of terms such as are used by Lammson and his co-workers[11] and by Gordon.[3, 4]

This strategy requires a predetermined terminology and order which will vary with the physician compiling the dictionary. The creation of the dictionary is itself a formidable task.

Combination techniques which use both coding and word search have been employed by Eiduson.[2] These input techniques include entry of data in structured form, but have severe limitations in that the capture

FIGURE II. FORMAT OF ORTHOPEDIC SURGICAL REPORT

¼
(OSS)
IDENTIFICATION NO.
BELLEVUE HOSPITAL — ORTHOPEDIC SURGERY — SURGICAL REPORT
NAME — SEX —
DATE OF SURGERY —
LOCATION —
BIRTH DATE —
AGE — RACE —
ATTENDING —
OPERATOR —
##
(ORTHOPEDIC PROCEDURE) —
(PREOPERATIVE DIAGNOSIS) —
(OPERATIVE DIAGNOSIS) —
(ANESTHESIA) —
(POSITION OF PATIENT) —
(TOURNIQUET) —
(INCISION) —
(APPROACH) —
(PROCEDURE) —
(FIXATION DEVICES) —
(CLOSURE) —
(SUCTION) —
(ANTIBIOTICS) —
(OTHER DRUGS) —
(CAST) —
(BLOOD) —
(PATHOLOGY) —
(STATUS POSTOPERATIVE) —
(KEY DATA) —
(DICTATED) —
++++

of data by the physician and secretarial staff is costly in time and labor.

Mention could be made of methods which attempt syntactical analysis of unstructured narrative data. The task is difficult, if not impossible, when ordinary English is used. The narrative material found in medical documents is a sublanguage in itself and each specialist has his own terminology and usage. Some constraints on the input must be used to permit successful study of the information contained in any medical document.

The Variable-Field-Length (V-F-L) format is a method of structuring medical information.[5, 8, 10, 13]

The foundation of the V-F-L format technique herein described is based on the observation that clinical information generated by the patient and collected by the physician usually follows a well-defined sequential pattern. Our goal is that the information obtained by the physician will be available for routine use in a form which can be stored, retrieved, and analyzed by the computer. Further, the physician should not be required to alter his usual pattern of work substantially. Therefore the method used in accumulating data for computer input should be a by-product of established recording procedures while placing a maximum amount of information in a computer-accessible form.

As a result of the development of this technique, the physician in any field can create unique structures, called *formats*, which obey certain specified constraints that permit the capture of [clinical] identification and narrative data. This work has been done with the departments of neurology, neuroradiology, neurosurgery, pediatrics, pathology, and others, and lends itself well to the field of orthopedic surgery. Previous demonstration of this technique (Lusskin et al.[12]) was limited to discharge reports and surgical reports. More powerful computer tools have now been made available (Univac 1108) and the types of orthopedic documents have been increased to include outpatient and follow-up information.

The Orthopedic Documents

Formats. Figures 1 through 5 present the formats of the five orthopedic documents presently in use. These formats consist of a fixed field and a variable field.

The fixed field contains identification information. The variable field consists of the body of the report and is presented in narrative form. Two orders of paragraphs are identified in the format. These are the

FIGURE III. FORMAT OF INITIAL ORTHOPEDIC ASSESSMENT OUTPATIENT REPORT

¼
(OAR)
IDENTIFICATION NO.—
BELLEVUE HOSPITAL — INITIAL ORTHOPEDIC ASSESSMENT
NAME— SEX—
DATE OF VISIT—
LOCATION—
BIRTH DATE—
##
(INFORMANT)—
(HISTORY)
 ORTHOPEDIC PROBLEM—
 MEDICAL—
(EXAMINATION)—
 GENERAL—
 ORTHOPEDIC—
 NEUROLOGICAL—
(DIAGNOSTIC STUDIES AVAILABLE)
(IMPRESSION)
 PRIMARY—
 SECONDARY—
 SYMPTOMS AND SIGNS—
 LOCALIZATION—
 ETIOLOGY—
(MANAGEMENT)
 DIAGNOSTIC STUDIES—
 TREATMENT—
 SURGICAL PLAN—
 CONSULTATION—
 REFERRAL—
(KEY DATA)—
(RETURN APPOINTMENT)—To Orthopedics—
(DOCTOR)—
++++

first- and second-order paragraphs. They order the data for computer processing and are identified specifically on the magnetic tape which is eventually produced. They closely resemble the usual headings of a standard discharge, operative, or other report.

We have added some coding to the narrative documents. In the

FIGURE IV. FORMAT OF ORTHOPEDIC INTERVAL VISIT REPORT

¼
(OIV)
IDENTIFICATION NO. —
BELLEVUE HOSPITAL — ORTHOPEDIC INTERVAL VISIT
NAME — SEX —
DATE OF VISIT —
LOCATION —
BIRTH DATE —
##
(INTERVAL HISTORY) —
(EXAMINATION) —
(IMPRESSION) —
(MANAGEMENT) —
(RETURN APPOINTMENT) — To Orthopedics —
(DOCTOR) —
++++

Orthopedic Discharge Summary (ODS) a second-order paragraph under "Hospital Course" labeled "Complications Code" is seen. This permits a brief three-letter word to identify the system which suffered a complication and permits easy search for complications. The instructions for this item are contained in the dictating instructions.

Similar is the Orthopedic Follow-up Report (OFR). Three paragraphs have been designated for one-word follow-up evaluation in addition to any narrative description of the patient's condition. The dictating instructions indicate the use of these paragraphs.

Figure 6 demonstrates a complete orthopedic discharge summary, with the identification and narrative data entered into the document.

Dictating Instructions

Specific dictating instructions must be prepared for the physicians who prepare the reports. These instructions are not complicated and can be mastered with little additional effort. They must be followed carefully or errors will appear in the final documents.

There are general and specific instructions prepared for each docu-

FIGURE V. FORMAT OF ORTHOPEDIC FOLLOW-UP REPORT

¼
(OFR)
IDENTIFICATION NO.—
BELLEVUE HOSPITAL — ORTHOPEDIC FOLLOW-UP REPORT
NAME— SEX—
DATE OF VISIT—
LOCATION—
BIRTH DATE—
##
(INTERVAL HISTORY)—
(EXAMINATION)
GENERAL—
ORTHOPEDIC—
NEUROLOGICAL—
(IMPRESSION)
PRIMARY—
SECONDARY—
SUBJECTIVE STATUS—
OBJECTIVE STATUS—
ECONOMIC STATUS—
(MANAGEMENT)
DIAGNOSTIC STUDIES—
TREATMENT—
SURGICAL PLAN—
(KEY DATA)—
(RETURN APPOINTMENT)—To Orthopedics—
(DOCTOR)—
++++

ment. The three outpatient documents (OIV, OAR, and OFR) are at present written out by the physician and the instructions appear in each of the prepared sheets upon which the house officer writes.

The instructions for dictation are available from this department and are included in *A Computerized Medical Record System*, Julius Korein, M.D., editor.[9]

Capture. Capture is the term used to describe the process by which physician-generated information is converted into signals within the ordering or structuring device—the format.

Each V-F-L format is designed to order a specific type of medical

FIGURE VI. COMPLETED ORTHOPEDIC DISCHARGE SUMMARY WITH IDENTIFICATION AND NARRATIVE DATA ENTERED INTO DOCUMENT

¼
(ODS)
IDENTIFICATION NO. — B303614
BELLEVUE HOSPITAL — ORTHOPEDIC SURGERY — DISCHARGE SUMMARY
NAME — Patrick SEX — Male
DATE OF ADMISSION — 05/24/71 DATE OF DISCHARGE — 06/12/71
LOCATION — 1-6
BIRTH DATE — 06/03/36
AGE — 35 RACE — White
SURGERY — None
PM — NA
##
(HISTORY)
CHIEF COMPLAINT — Acute low back pain radiating to the right buttock secondary to concrete shoveling since 05/24/71.
PRESENT ILLNESS — The patient denies pain radiating down either leg. There is no past history of low back pain. The pain is not exacerbated by coughing. Pain relieved by bed rest on firm mattress. Unable to bend back more than 20 degrees without severe pain in any direction.
PAST HISTORY — Patient has had no medical problems with his head, eyes, ears, nose or throat. No previous cardiovascular problems. No gastrointestinal history. No other medical history.
FAMILY HISTORY — NA.
(PHYSICAL)
VITAL SIGNS — BP 120/80. P 80. RR 12. T 99.
GENERAL PHYSICAL — Well developed, well nourished male in no acute distress. Head, eyes, ears, nose and throat within normal limits. Neck thyroid not palpable. Chest clear to percussion and auscultation. Abdomen soft. No organomegally noted.
ORTHOPEDIC EXAMINATION — Upper extremities completely within normal limits. Spine revealed marked paravertebral muscle spasm on the left and right side of the low back was noted at time of admission. Straight leg raising positive on left side at approximately 25 degrees and positive on the right side at approximately 20 degrees. FABER test negative. Jugular compression test negative.
NEUROLOGICAL EXAMINATION — Alert and oriented. Cranial nerves grossly intact. Gait not tested because of acute low back pain. No ataxia. Reflexes normal. No sensory deficit.
(LABORATORY FINDINGS)
URINE — 05/24/71, normal.
BLOOD — 05/24/71, WBC 6200.
X-RAYS — 05/24/71, PA lateral and oblique reveals straightening of the lumbar spine. No spondylosis. No spondylosthesis.
OTHER LAB — NA.
(COURSE)
COURSE IN HOSPITAL — The patient was treated with Valium and bed rest and narcotic analgesics for 48 hours. The patient responded well to bed rest and was discharged in satisfactory condition on 06/12/71.
COMPLICATIONS CODE — None.
STATUS AT DISCHARGE — Improved.
DISPOSITION — Home.
(DIAGNOSIS)
ADMISSION DIAGNOSIS — Lumbar sacral sprain.
FINAL DIAGNOSIS — Lumbar sacral sprain.
SECONDARY DIAGNOSIS — None.
(KEY DATA) — None.
(DICTATED) — H. Goodman/ps, 06/14/71.
++++

information, such as the discharge summary, operative report, or follow-up document. Any number of different formats can be created.

Each format consists of a fixed field for identification data and a variable field for narrative clinical data. The orthopedic formats presently created are described under Orthopedic Documents.

The format orders the data and permits identification of all portions of the document. It is followed by the physician and is used by the typist with her keyboard equipment to generate signals for the computer.

The physician collects information from and about the patient. In the computerized record-keeping system, as in standard usage, he records this information with dictating equipment or by working on prepared forms. The sequence of recording is determined by the format which is used for each medical document.

The dictated or written information is then processed by a secretary or typist trained in techniques of data capture. Two methods are now being used to capture the medical information. The first is the production of magnetic or paper tapes by means of special typewriters. The second is the use of teletype terminals to enter data directly into the computer.

The production of magnetic or paper tape is an off-line process. "Off-line" means that the computer plays no role in the conversion from keyboard device to computer readable form. This is relatively inexpensive compared to on-line techniques. Most of our documents are so handled.

The direct capture of information under the control of the computer is called on-line data capture. We are currently using teletype terminals.

Each method permits the production of paper reports—hard copy—which can then be edited and placed in the presently maintained paper records of patients and in the service file and other files.

Off-Line Capture

The data-processing secretary types the dictated or written report on a magnetic tape (IBM-MT/ST) or paper tape (Dura or Frieden) typewriter, simultaneously producing the hard copy and the signal-containing tape—the document tape.

Each of these typewriters permits previously created master format tapes to drive the equipment. The format tape adds that information which is constant for a particular document—i.e., the identification headings and paragraph headings—to the hard copy and output document tape. The secretary adds the physician's written or dictated information. Also added to the document tape are nonvisible keyboard symbols which identify it further. Invisible symbols are built into the structure of the format to identify the different types of paragraph. The typed hard copy of the document (format plus content) is returned to the physician for final editing. It is then retyped when necessary.

To reiterate: the first step is the collection of information about the patient by the physician. The second is the recording of the information and its transmittal to the secretary. The third step is the formating of the data by the format tape which designates its structure on paper tape or magnetic tape.

Changes of technology need in no way modify the original formats, nor would they lead to loss of completed documents. Conversion from one type of machine to another is possible and should not concern the physician.

The second method of data capture involves the use of on-line computer terminals. A teletype terminal is used to introduce the narrative document directly into the computer. The document, usually an outpatient document, is put into the computer, and its identification section is filled in from a previously created file. The document is then processed by the computer and placed in temporary storage on the magnetic drum memory in a form that is directly analogous to the paper or magnetic tape described in the previous section.

Input and Storage

Because of the equipment available to us we must convert the initial signal-containing document tapes produced by the MT/ST or paper tape typewriters.

Paper-tape conversion. At the time of writing the conversion of the punch-paper tape, whether the product of the Dura machine or of the Friden Flexowriter, is converted on a Univac 9300 which contains a tape reader. This generates a magnetic computer-compatible tape. We are planning to use a PDP 8, which is connected to the Univac 1108 by communication lines. The Univac 1108 is the large computer used in this system.

Once the paper tapes have been converted to magnetic tape their data are ready to be entered into the computer as input data.

MT/ST/tapes. Once typed on, the small reels of MT/ST tape, called cartridges, contain magnetic signals which are not computer-compatible. These cartridges are converted by a Digi-Data and transferred onto somewhat larger reels called minireels which are computer-compatible. Each minireel can contain approximately 170 pages of document. The MT/ST cartridge can hold 15 to 20 single-page documents.

As the Digi-Data converts MT/ST cartridges, it produces a typed hard copy. When this copy is checked by the operator and found to be without conversion error, the minireel is boxed, labeled, and readied for its trip to the central computer complex. The reel number assigned to this tape is logged and the corresponding number of documents are logged.

The data collector delivers the minireel containing digitized documents to the Univac 1108. Then two 10-inch computer tapes are assigned to receive the data from the minireel. The first will have the data read directly onto it; the second will contain the data rearranged by a computer program called MINDEX.

MINDEX

MINDEX is the computer program which identifies and validates the format of each document and which converts the original structure of the documents into a structure which can be searched and analyzed more easily. It does this by constructing a directory of the paragraphs which locates the beginning of each paragraph.

The purpose of this operation is to place the narrative content in a form which will simplify analysis and enable the user to examine the content of each paragraph more easily.

MINDEX validates the documents through comparison with the format library. If the document is found acceptable to MINDEX, it is assigned a unique document number and is then stored. At the end of a MINDEX run, the last document number is recorded. If an error is discovered in the format while the document is in MINDEX, the document which contains the error is rejected. When the entire tape has been run and a listing is made of all the documents that have been MINDEXed, the documents containing errors are printed out and the error is identified in order to facilitate the editing of the documents.

As of the time of writing there is no editing program to correct documents once they are within a computer magnetic tape.

The document number which MINDEX assigns to each accepted document is recorded by the secretary. The secretary-typist then checks all MINDEX rejects (errors) and takes the appropriate steps for reentering these errors in a corrected version. The MINDEX reject is a printout of either the entire document or of a portion of that document, specifying the particular error. The typist need transfer only the incorrect document (in the case of IBM MT/ST captured data) from its original cartridge, inserting the corrections as the transfer is made.

The method for correcting documents is detailed in the IBM MT/ST manual. After all errors have been corrected and successfully entered through MINDEX, the secretary-typist can erase the MT/ST cartridges and use them again. The punch-paper tape can also be destroyed after corrections and recopying have been completed.

This completes the first stage of the operation. We now have two magnetic tapes. These store our data and permit the information to be manipulated by computer.

On-line captured information. Documents captured through computer terminals and temporarily stored on the magnetic drum memory of the computer are processed by MINDEX in the same way as the off-line information. Ten-inch computer tapes are also produced by this procedure. The on-line method permits immediate checking of the formats and assists the typist in entering the data but is more expensive.

Figure 7 summarizes the varied input, conversion, and storage routes leading to MINDEX and to retrieval and analytical operations.

Data Management and Quality Control

While recording of the documents entered into the system and control of the operation can be fairly simple when only one service is using the system, major problems will develop when a large facility begins to process all or a large part of the medical information generated. These problems are controlled by supervision and logging.

The logging procedure for a computerized medical-records system is most important. It should serve as a central file from which a user can readily discover the location of desired information. A logging system should be so up to date as to eliminate document duplication or

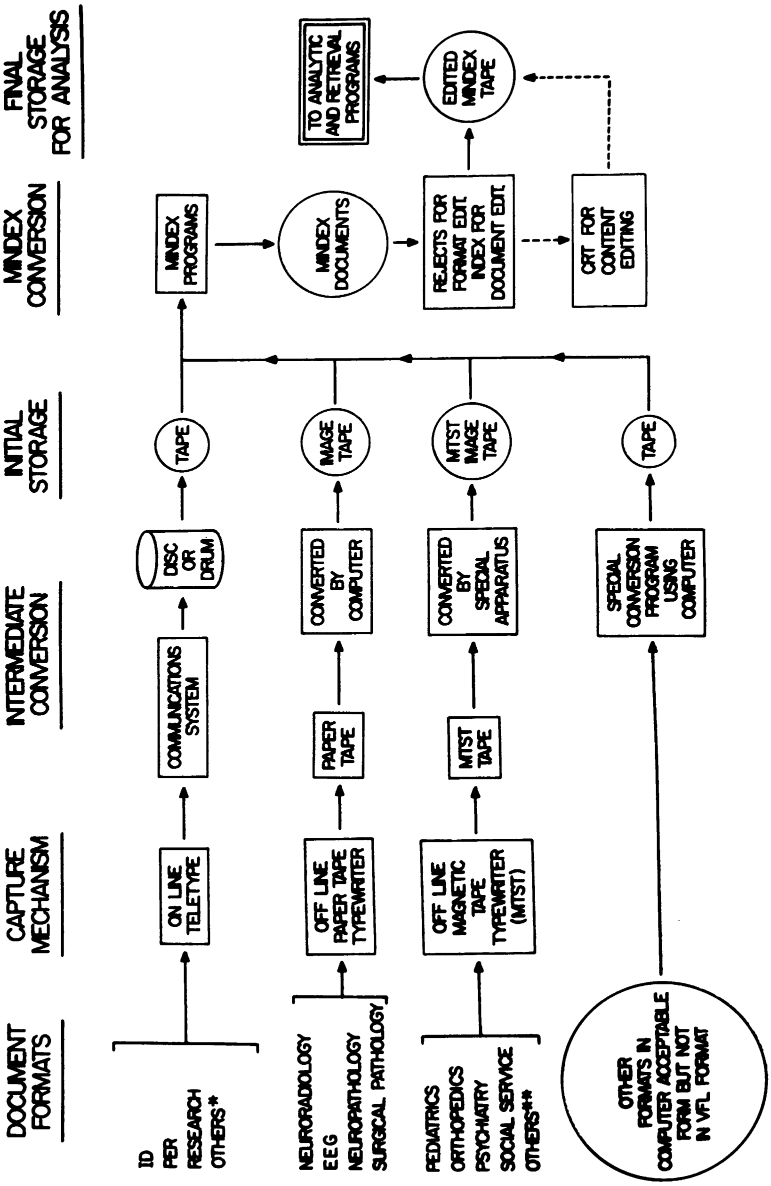

Fig. 7. Capture techniques using VFL format and information flow to storage in MINDEX form. Reproduced by permission from Anderson, J. and Forsythe, J. M., editors: *Information Processing of Medical Records*. Amsterdam, North-Holland, 1971.

the loss of documents. It should identify the computerized data in the system at any time. A properly organized and managed log will also insure optimum speed in correction and retrieval of off-line documents.

Supervision

We have attempted to maintain as much supervision over the multiple, disparate parts of the system as is needed to produce a reliable record of the patient. A systems analyst has been assigned to this function. It is his responsibility to follow and verify the daily flow into the system as well as to check the subsequent output against that initial output, and thus detect and correct errors as they occur. This data-management function should be sufficiently flexible to be able to meet the constantly increasing volume of data being processed by the computer. The systems analyst is the liaison between the computer staff (programmers, mathematicians, and others), the medical staff, health professionals, and the secretarial and typing staff. The need for this activity is indicated by the utilization of these services by a variety of persons. The interacting groups of health professionals, computer personnel and secretarial personnel must obtain correct information in response to their system-oriented questions.

Analysis

Although it is not the purpose of this paper to discuss analytic techniques in detail, there is some value in outlining the strategy by which the narrative analysis may be performed. The basic assumption underlying strategy is that medically meaningful analysis may be made by determining the presence or absence of words, their synonyms, and modifiers in certain places in a document. No syntactical analyses are made to define meaning. The only exception to this statement is that the user may request that a combination of words being searched out occur in some easily delimited grammatical subunit: e.g., a sentence or paragraph. To obtain a good analysis, successful and meaningful, under these circumstances, it is necessary for the physician to plan carefully and to play an intensive role in the search.

Since word combinations in various locations in a document may have differing significance in differing specialties, the specification for each analysis must be tailor-made. For different studies the lists of words which may be considered synonymous may be different.

As the task of determining lists of synonymous words is not trivial, we may use the computer to assist in this operation by making it produce word-frequency lists from a representative or complete set of the type of documents to be analyzed.

For example, the search for "osteomyelitis" would include a set of synonymous terms such as bone infection, bone abscess, osteitis, bone sepsis, subperiosteal abscess, septic necrosis, etc., and would specifically exclude terms deemed not to imply bone infection: i.e., bone syphilis, lues, aseptic necrosis, radiation necrosis, etc.

In actual practice, the review of a word list which includes all misspellings and mistypings is simple. Iteration after review of the material produced after synword search leads to further refinement of both synonyms and excluded words.

Processing a Digidata minireel through the computer for MINDEX is rapid and costs about $5.00. To search a MINDEX 10-inch reel and produce a word list takes two minutes of computer time at a cost of $20.00 to $30.00.

Problems With Input

Problems that should be considered in relation to input include: 1) general acceptance of the method by the physician and typists, 2) editing and correcting, 3) conversion of typewriter paper tape to computer magnetic tape, and 4) logic.

Initially some resistance by the physician in dictating in the prescribed manner was encountered. However, this was overcome when the physician realized that the sequence followed the usual pattern of dictation and that his own service modified the documents for its own specified purposes. The method has been used in routine dictating and typing of the discharge summaries of the Neurology and Neurosurgery Service of New York University-Bellevue Medical Center since February 1963, and has been extended to neuroradiology reports in March 1964 and to electroencephalogram reports and research protocols since January 1965. A study of head injury, initiated in 1964, has demonstrated the utility of this method for clinical research. The dictation of neuroradiology and electroencephalogram reports was accepted immediately and took no longer than the dictation of previous reports. The secretary's typing time may have been increased initially, but this has depended upon her experience with the special typewriter, the number of errors made, and the type of machine used. Newer punch-paper-tape typewriters allow a faster rate of typing than the ordinary electric typewriters. The initial resistance to the technique appears to be related to unfamiliarity; once this is overcome, little objection is noted. The orthopedic house staff has adapted well to each addition to the system. Routine application of the method has been used to type and enter into the computer more than 50,000 clinical documents which have been dictated by physicians.

As we have noted, the problem of editing entails some difficulty in that this must be done carefully and requires a correction of the intitial typewriter tape.

The transfer of the paper tape to magnetic tape at present is performed in groups of 100 to 200 paper tapes, each tape representing a case summary or report. This takes about one hour. The IBM cartridges are converted in batches of tens. The possibility of transferring the data over communication lines from the source to the computer is known to be feasible. We are presently considering the use of a microwave transmitter for the transfer of data.

Another difficulty involves problems in posing well-defined questions by the physician. On occasion there were marked differences of interpretation of a single concept among the physicians dictating the reports. The computer will do exactly what we tell it to do, no more, no less. If we have difficulty defining the problem, the errors and ambiguities made by the machine increase. Finer and more explicit delineation and use of medical terminology may be another useful by-product of the use of the computer by the physician.

Conclusion

The application of computers to orthopedic narrative data for the purpose of storage and retrieval is now feasible and has been discussed in terms of a specific method of input which utilizes the V-F-L format and the punch-paper typewriter or magnetic-tape typewriter. The progress thus far is encouraging and appears to point the way to the use of modern technology toward reasonable solutions of problems of medical record-keeping and of clinical research with narrative data. The specific formats which have been created for the orthopedic documents described may be used to create a computer-manageable bank of clinical data which can be made available for storage and for retrieval on demand. In addition, analytic programs can be brought into action to correlate medical events from such a body of stored information.

Acknowledgment

We acknowledge the contributions that were made toward the development of this system by M. Lyman, M.D., associate professor of clinical pediatrics and associate program director, and by L. Tick, Ph.D., professor of medical statistics in the Department of Radiology, New York University School of Medicine, New York, N. Y.

REFERENCES

1. Brolin, I.: ***Current Studies in Data Processing of Radiological Reports.*** Rontgendiagnostisk avd. I, Goteborg SV, Sweden, Sahlgrenska sjukhuset, 1967-68.
2. Eiduson, B. T.: ***Psychiatric Case History Event System: Transcription Procedures with Lexicons.*** Los Angeles, Calif., Reiss-David Child Study Center, 1966.
3. Gordon, B. L.: Terminology and content of the medical record. *Proc. Conf. Information Systems:* 53-62. National Center for Health Services Research and Development, Dept. of HEW.
4. Gordon, B. L.: Regularization and stylization of medical records. *J.A.M.A. 212:*1502-07, 1970.
5. Korein, J., Tick, L., Woodbury, M. A., Cady, L. D., Goodgold, A. L. and Randt, C. T.: Computer processing of medical data by Variable-Field-Length format. *J.A.M.A. 186:*132, 1963.
6. Korein, J., Goodgold, A. L. and Randt, C. T.: Computer processing of medical data by Variable-Field-Length format: II. Progress and application to narrative documents. *J.A.M.A. 196:*950, 1966.
7. Korein, J., Bender, A. L., Rothenberg, D. and Tick, L.: Computer processing of medical data by Variable-Field-Length format: III. Statistical analysis of narrative content. *J.A.M.A. 196:*957, 1966.
8. Korein, J., Goodgold, A. and Randt, C. T.: Computer processing of narrative medical data: A critical evaluation of results in 100 cases of head injuries. *Neurology 14:*202, 1966.
9. Korein, J.: ***A Computerized Medical Record System,*** vols. 1 and 2. Dept. Health, Education, and Welfare, Public Health Service. Nat. Center for Health Services, Research, and Development, Washington, D.C. In press.
10. Kricheff, I. I., Chase, N. E. and Korein, J.: Computer processing of clinical medical data. Computers in radiology. *Proc. Int. Meeting,* Brussels, 1969. New York, Karger, 1970.
11. Lamson, B. G.: *Computer Assisted Data Processing in Laboratory Medicine. Computers in Biomedical Research,* Stacy, R. W. and Waxman, B., editors. New York, Acad. Press, 1965, vol. 1, p. 353.
12. Lusskin, R., Thompson, W. A. L., Korein, J. and Krause, E.: Computer handling of orthopedic clinical information: Collection, storage, retrieval. Proc. Amer. Academy Orthopaedic Surgeons. *J. Bone Joint Surg., 40-A:*1267, 1968.
13. Robinson, R. E., III.: Acquisition and analysis of narrative medical data. *Proc. Conf. Med. Information Systems.* National Center for Health Services, Research and Development, Department of HEW, Rockville, Md., January 1970, pp. 111-126.
14. Seed, J.: Restricted data formats: Use of data mechanization and computers in clinical medicine. *Ann. N.Y. Acad. Sci. 161:*484, 1969.