Lynette Hirschman,* Naomi Sager,* and Margaret Lyman, M.D.[†]

*Linguistic String Project, Faculty of Arts and Science †Department of Pediatrics, NYU Medical Center New York University, New York, New York

Summary - This paper describes an experimental computer program for the application of health care review criteria to hospital discharge summaries. The use of the computer in this process would make it possible to speed up the routine screening of patient records; it could also facilitate experimental evaluation of alternate proposed audit criteria. The computer program has two components. The first component creates a structured form of the information contained in natural language medical records. It maps the words of each sentence into labelled columns of a table (or information format) according to the type of medical information contained in each word. This structured information is suitable for use as a data base in many areas of clinical research. The second component consists of a set of retrieval routines, each of which corresponds to a criterion of the health care evaluation form, e.g., was the patient afebrile at discharge? The retrieval component is built up in modular fashion, so that basic routines can be used in other applications. The application of this program to a sample hospital discharge summary is presented and compared to the results obtained by a physician reviewer.

Introduction

Review procedures to assess the quality of medical, or health, care have been a requirement since 1967 when the Medicare program included medical care evaluation studies -- one type of review procedure -- in its utilization review activities. With passage of the 1972 amendments to the Social Security Act, a Bureau of Quality Assurance was established within the Department of Health, Education and Welfare, and the Professional Standards Review Organization program (PSRO) came into being. The PSRO program is responsible for either doing the review itself or seeing that procedures to assess the quality of medical care given to certain patients are carried out by all health care providers. Medicare, Medicaid, and certain other government beneficiaries must have their health care reviewed. PSRO's are also encouraged to review care given to patients covered by other third party payers. All major proposals for National Health Insurance include requirements that PSRO's review hospital, nursing home and ambulatory care.

Additional impetus for review has come from the Joint Commission for the Accreditation of Hospitals (JCAH). The JCAH has included, among its many requirements, systematic review of utilization of hospital services and medical care evaluation studies. Thus, two essential "boards" require review of medical care -- government funding agencies and hospital accreditation.

Now that medical care has come under regular review, the instruments used in the review process are receiving widespread attention. One of the earliest systems made available to all hospitals was the Performance Evaluation Procedure (PEP) of the JCAH. The PEP system for medical care evaluation endeavors to measure medical care given against objective, predetermined explicit criteria ("medical audit"). In the JCAH system, hospitals select problem areas, e.g., diagnoses, procedures (especially surgical), and treatments; the experts formulate the criteria to be applied against randomly selected patient records; nurses or trained record reviewers match the criteria against the information in the records and tally those records which meet or vary from the criteria. When variations are significant, medical educators and administrators are called upon to correct the practices or provide whatever is needed to correct the deficits. At a later date, a re-audit is performed to ensure that correction has taken place.

The criteria used in medical audit are not universally agreed upon. All are agreed that medical audit of an individual patient/physician may not always assess the quality of medical care and that the human effort involved is enormous. A critical need exists for assessing the value of the procedures and criteria used, both in demonstrating ability to measure quality care and in pointing to problems which may develop in the distant future as a consequence of the care.^{1,2}

The work to be described here, that of automating the application of review criteria to patient records, offers a number of potential advantages: 1) it performs a screen of patient records in advance of the record reviewer to find particular criteria not met and therefore in need of human review, thus reducing the work load of the reviewer; 2) it enables one to extend or change the criteria without imposing a second round of human review of the same or different records; 3) it provides a body of information about medical care that could be subjected to a variety of analyses as new instruments for determination of quality assurance are developed.

Document Processing

The computer program for health care evaluation

(PDS) IDENTIFICATION NO -HOSPITAL-PEDIATRIC DISCHARGE SUMMARY SEX - F NAME -DATE OF ADMISSION - 1/10/72 DATE OF DISCHARGE -1/12/72 LOCATION -BIRTH DATE -## DOCUMENT NUMBER (REFERRING PHYSICIAN) - NONE GIVEN. (REASON FOR ADMISSION) -SWOLLEN, PAINFUL HANDS. VOMITING. SYMPTOMS OF 18 HOURS DURATION. (PERTINENT HISTORY) -HOSPITAL ADMISSION. PATIENT HAS SICKLE CELL DISEASE. WAS WELL UNTIL PRESENT ILLNESS - THIRD 10 HOURS BEFORE ADMISSION. PATIENT BEGAN TO VOMIT. THEN SHE DEVELOPED PAINFUL HANDS. TEMP ELEVATION TO 101 DEGREES WAS NOTED. SHE WAS CONVALESCENT FROM CHICKEN POX AT A VISIT 1 WEEK BEFORE ADMISSION. SIGNIFICANT PAST HISTORY - TWO ADMISSIONS FOR MENINGITIS. 1 TRANSFUSION REQUIRED AT EACH ADMISSION. HAS BEEN TAKING FOLIC ACID DAILY. PENICILLIN GIVEN WHEN THERE IS EVIDENCE OF INFECTION. (EXAMINATION ON ADMISSION) -TMP 99.6, PU 120, RR 16, WEIGHT 19.5 LBS. WELL DEVELOPED, WELL NOURISHED. OCCASIONALLY HOLDS AND RUBS HANDS AND FOREARMS. NO MENINGISMUS OR ABNORMAL NEUROLOGIC FINDINGS. DORSAL SURFACE AND PROXIMAL PALMAR SURFACES OF BOTH HANDS ARE SWOLLEN, WARM AND PAINFUL. (IMPRESSION ON ADMISSION) -HAND-FOOT SYNDROME WITH SICKLE CELL DISEASE. (COURSE IN HOSPITAL) -ADEQUATE HYDRATION WAS ACHIEVED WITH I.V. FLUIDS AND CLEAR LIQUIDS ORALLY. ON THE DAY AFTER ADMISSION THE PATIENT TOOK A REGULAR DIET. HER HANDS REMAINED SOMEWHAT WARM AND SWOLLEN; HOWEVER PAIN SEEMED TO SUBSIDE. (STATUS AT DISCHARGE) -IMPROVED. SLIGHT RESIDUAL SWELLING OF BOTH HANDS. (LABORATORY DATA) -HEMATOLOGY - HEMATOCRIT 28; SED RATE 27; WBC 23,700; RETICS 29.8%. NEXT DAY WOC 16,500 WITH 36% POLYS, 58% LYMPHS. URINE - URINALYSIS NORMAL. BLOOD CHEMISTRIES - ELECTROLYTES NORMAL. BACTERIOLOGY - 3 BLOOD CULTURES NEGATIVE. OTHER - CHEST XRAY NORMAL. STOOL GUIAIC NEGATIVE. (DIAGNOSES) -HAND-FOOT SYNDROME. SICKLE CELL DISEASE. (PLAN AT DISCHARGE) -NO MEDICATIONS REQUIRED. TO BE SEEN FOR CHECKUP IN 2 DAYS. TO BE FOLLOWED IN HEMATOLOGY CLINIC. (RETURN APPOINTMENT) -HEMATOLOGY 01/18/72. (ABSTRACT) -PATIENT, 1 YEAR OLD, IS KNOWN TO HAVE SICKLE CELL DISEASE AND 2 EPISODES OF MENINGITIS. DEVELOPED SWOLLEN, PAINFUL AND WARM HANDS. HAD SEVERAL EPISODES OF VOMITING PRIOR TO ADMISSION. LABORATORY STUDIES DID NOT REVEAL ANEMIA OR SYSTEMIC INFECTION. HYDRATION THERAPY AND BED REST WERE PROVIDED, WITH IMPROVEMENT IN 48 HOURS. SHE WAS DISCHARGED IMPROVED. TO BE FOLLOWED IN HEMATOLOGY CLINIC. (DOCTOR) -/SPECIAL STUDY/FDH. DATE - 07/01/72. ++++

FIGURE 1

consists of two distinct components, one which creates the data base, the other which retrieves the desired information from it. The creation of a data base from the sentences of the original medical record has been described in earlier publications.^{3,4} It is based on the use of a tabular information format whose columns correspond to the distinct kinds of information contained in the type of document being processed (in this case, hospital discharge summaries). Figure 1 shows a complete sample discharge summary of the type that we have processed; Table 1 shows a partial information format for this material, with entries generated from several sentences of text. The format contains columns for the relevant types of medical information, such as patient, medical treatment (MED), observed sign or symptom, part of body, etc. It also contains columns for various informational markers such as negation (NEG), uncertainty (MODAL), words indicating change of state (CHANGE), duration, etc. The full format also contains a number of columns for time information, not shown in Table 1.

The formatting component first performs a linguistic analysis of each input sentence in the document and records relevant grammatical information (for example, subject-verb-object relations and head-modifier relations in each sentence). Next, certain complex grammatical relations are converted into simpler relations by informationpreserving transformations. For example, hydration therapy and bed rest were provided is converted into two complete active assertions, () provided hydration therapy and () provided bed rest. Next, the formatting component maps the words of each input sentence into the format columns corresponding to the type of information contained in the word, based on the grammatical information already obtained and on lexical information obtained from a computer lexicon for medical vocabulary. A final stage of processing fills in implicit word occurrences that can be reconstructed from the context. For example, in the sentence Her hands remained somewhat warm and swollen; however pain seemed to subside., the pain referred to is in the hands, but this is not explicitly mentioned. In order to facilitate retrieval, the program normalizes the data table by filling in the missing words, as illustrated in line 7 of Table 1, where hand plural has been filled in, marked with square brackets to show that it did not occur in the original text.

Retrieval Routines

The second component of the computer program for health care evaluation consists of a set of retrieval routines corresponding to the criteria of a health care evaluation form. The form chosen for implementation is the health care evaluation form used in the hospital which provided the discharge summaries for this experiment. Figure 2 shows a set of evaluation criteria for hand-foot syndrome in sickle cell disease, along with the screening results for the document shown in Fig. 1. The figure contains two sets of results, those generated by computer and those filled in by a physician working from the same documents.

In the form used here, the criteria are divided into principal criteria (numbered) and alternate criteria (labelled with letters, as in 12A). The standard (abbreviated STD in Fig. 2) appears in the first column; it is the value expected for that criterion in a given document. A standard of 100% means that this information should be found in the document; for example, criterion 1: pain in hands and/or feet should be present for a patient diagnosed as having hand-foot syndrome. Conversely a standard of 0% means that the information should not be present, e.g., criterion 13: mortality. A document that does not meet the standard for one or more criteria (indicated by an entry in the VARiation column) is screened out for further examination and review by trained medical staff. This means that the computer is expected to perform only the initial routine screening of the documents; it is not expected to make subtle medical judgments.

Columns 2 through 4 of Fig. 2 show the computergenerated results for the document of Fig. 1. If the main criteria is met, a YES is printed in the column labelled ELement (see criteria 1 and 2 in Fig. 2). If an alternate criterion is met, a YES is printed in the column labelled EXception. However if neither the main criterion nor the alternate (if present) is met, then a YES is printed in the column labelled VAR, which records any variation from the expected standard of column 1 (see criteria 4-9 in the computer results of Fig. 2). Any entry in the VARiation column will cause the document to be screened out for further detailed review by medical staff. To the right of the computer-generated results are the results obtained independently by manual coding. Differences with the computer results are circled and will be discussed below.

To aid the review process for documents failing a criterion or a sub-part of a criterion, the computer program prints out a set of more detailed messages below the table itself (bottom of Fig. 2). For example, determination of the state of hydration can be recorded explicitly (e.g., <u>patient well</u> <u>hydrated</u>); or it can be determined implicitly by noting a number of physiological factors, as explained in the message for element 4 at the bottom of Fig. 2. These messages are issued where there are multiple subparts for a given criterion, to aid the reviewer in determining in detail which information is missing.

The retrieval sub-routines are developed in cooperation with a medical consultant who provides an elaboration of the criteria in those cases where the criteria do not correspond to the likely formulation of the information in a discharge summary. For example, criterion 4: determination of state of hydration is not likely to appear in the discharge summary as stated. The retrieval routine is formulated to look either for a description of the patient's state of hydration (e.g., patient well hydrated), or alternatively, for a number of physiological tests or observations indicating the patient's state of hydration: history of diarrhea or vomiting; urine specific gravity; skin turgor; moistness of mucous membrane; change in weight. In addition, the medical consultant furnishes lists of paraphrases and equivalent formulations (e.g., for pain in criterion 1, the routine accepts pain, painful, soreness, tenderness,...). The consultant also furnishes lists of medical classes, for

TABLE 1 PARTIAL INFORMATION FORMAT FOR TWO PARAGRAPHS FROM FIG. 1

TEXT

(COURSE IN HOSPITAL) -

ADEQUATE HYDRATION WAS ACHIEVED WITH I.V. FLUIDS AND CLEAR LIQUIDS ORALLY. ON THE DAY AFTER ADMISSION THE PATIENT TOOK A REGULAR DIET. HER HANDS REMAINDED SOMEWHAT WARM AND SWOLLEN; HOWEVER PAIN SEEMED TO SUBSIDE.

(STATUS AT DISCHARGE) -IMPROVED. SLIGHT RESIDUAL SWELLING OF BOTH HANDS.

INFORMATION FORMAT

	CODE	TTME*	PATTENT	MD-ACT	MED	BODY	BODY	NORM-	OUANT	SIGN-	NEC	MODAT	CHANCE	DURA-	END
1.	COPDS 3.1.1	HOSP		ACHIEVE	HYDRA- TION (ADE- QUATE)				20/11/1			FIORE		1104	
2.		HOSP			FLUID PLURAL (IV)										
3.		HOSP			LIQUID PLURAL (CLEAR) (ORALLY)										
4.	COPDS 3.1.2	HOSP	PATIENT (THE)	TAKE			DIET	REG- ULAR							
5.	COPDS 3.1.3	HOSP				HAND PLURAL (HER)			SOME- WHAT	WARM				RE- MAIN	
6.		HOSP				HAND PLURAL (HER)				SWOLLEN				RE- MAIN	
7.		HOSP				HAND PLURAL				PAIN		SEEM			SUB- SIDE
8.	STPDS 3.1.1	DISCH											IM- PROVED		
9.	STPDS 3.1.2	DISCH				HAND PLURAL (OF BOTH)			SLIGHT RESI- DUAL	SWELLING					

TIME: FOR COMPACTNESS, ONLY THE GENERAL TIME RELATIVE TO THE CURRENT HOSPITALIZATION IS SHOWN HERE. <u>HOSP</u> = DURING CURRENT HOSPITALIZATION; <u>DISCH</u> = AT DISCHARGE. THE FULL FORMAT CONTAINS COMPLETE TIME INFORMATION (E.G. <u>ON THE DAY AFTER ADMISSION</u> IN COPDS 3.1.2).

MATERIAL IN SQUARE BRACKETS IS FILLED IN FROM IMPLICIT CONTEXTUAL INFORMATION, DURING THE FINAL STAGE OF DATA BASE GENERATION.

FIGURE 2

PEP CRITERIA - HAND-FOOT SYNDROME IN SICKLE CELL DISEASE

	COMPUTER RESULTS			HAND-CODED RESULTS*				
	STD	EL	EX	VAR	EL	EX	VAR	
(DIAGNOSIS OF HAND-FOOT SYNDROME)								
(1 PAIN IN HANDS AND/OR FEET OR EQUIVALENT)	100%	YES	0		I ∕			
(2 SWELLING OF HANDS AND/OR FEET)	100%	YES	0		/			
(DIAGNOSIS OF SICKLE CELL DISEASE)								
(3 ONE OF FOLLOWING)	100%	YES	0		√			
(3 (1) POSITIVE SICKLE CELL PREPARATION)		0	0					
(3 (2) HEMOGLOBIN ELECTROPHORESIS = HGS)		0	0					
(3 (3) STATEMENT IN HISTORY 'KNOWN SICKLER' OR EQUIVALENT)		0	YES			/		
(SPECIAL PROCEDURES AT ADMISSION)								
(4 DETERMINATION OF STATE OF HYDRATION)	100%	0	0	(YES)	\bigcirc			
(5 HEMATOLOGIC STATUS)	100%	0	0	YES	\bigcirc			
(6 X-RAYS OF HANDS AND/OR FEET)	100%	0	0	YES	Ŭ		\checkmark	
(7 BLOOD CULTURE)	100%	0	0	(YES)	\bigcirc			
(DISCHARGE STATUS)				\mathbb{X}				
(8 REDUCTION OR DISAPPEARANCE OF PAIN AND SWELLING)	100%	0	0	(YES)	\oslash			
(9 AFEBRILE)	100%	0	0	YES			1	
(10 STABLE HEMATOCRIT OR HEMOGLOBIN)	100%	0	0	YES				
(11 PLAN FOR ONGOING CARE)	100%	YES	0		\checkmark			
(12 LENGTH OF STAY - 2 - 8 DAYS)	100%	YES	0		1			
(12 A TRANSFERRED OR SIGNED OUT OR DIED)		0	0					
(12 B COMPLICATIONS OF SICKLE CELL DISEASE)		0	0					
(13 MORTALITY)	0%	YES	0		\checkmark			
(COMPLICATIONS OF SICKLE CELL DISEASE)								
(14 APLASTIC CRISIS)	0%	YES	0		\checkmark			
(14 CM1 = MONITORING HEMATOCRIT/HEMOGLOBIN)		0	0					
(14 CM2 = TRANSFUSIONS)		0	0					
(15 PNEUMONIA)	0%	YES	0		1			
(15 CM = CONFIRMED BY CHEST XRAY)		0	0					
(16 MENINGITIS)	0%	YES	0		\checkmark			
(16 CM = CSF POSITIVE CULTURE)		0	0					
(17 OSTEOMYELITIS)	0%	YES	0		\checkmark			
(17 CM = PROVEN BY X-RAY)		0	0					

EL 4 (DETERMINATION OF STATE OF HYDRATION) REQUIRES EITHER STATEMENT OF HYDRATION/DEHYDRATION OR THREE OF THE FOLLOWING -

1 HISTORY OF VOMITING OR DIARRHEA

- 2 URINE SPECIFIC GRAVITY
- 3 SKIN TURGOR
- 4 MUCOUS MEMBRANE
- 5 CHANGE IN WEIGHT
- CONDITIONS 1 WERE FOUND TO BE TRUE

EL 5 (HEMATOLOGIC STATUS) REQUIRES ALL OF THE FOLLOWING -

- 1 WHITE BLOOD COUNT
- 2 HEMATOCRIT
- 3 SED RATE
- 4 RETICULOCYTE

NONE OF THE FOUR CONDITIONS WAS FOUND

FOR EL 8 ONLY REDUCTION/DISAPPEARANCE OF PAIN WAS NOTED

* Differences between computer-generated and hand-coded results are circled.

example, criterion 1 calls for finding pain in hands and/or feet or equivalent; the sub-routine looks for a BODY-PART belonging to the set of EXTREMITY-WORDs, as provided by the medical consultant, namely <u>hand</u>, <u>foot</u>, <u>leg</u>, <u>extremity</u>, <u>finger</u>, <u>toe</u>, etc.

Results

We can use Fig. 2 to compare the computergenerated results with those obtained by the physician reviewer. We note that 13 out of the 17 answers are in agreement. The computer-generated and the hand-coded results differ in criteria 4, 5, 7 and 8.

Criterion 4: Determination of state of hydration (under <u>special procedures at admission</u>). Inspection of the original document (Fig. 1) shows no explicit mention of state of hydration at admission, and only one of the required implicit factors (vomiting). There is one mention of hydration, namely in the first sentence of the COURSE paragraph (Adequate hydration was achieved with I.V. fluids and clear liquids orally); however this sentence does not tell us about the patient's state of hydration at admission. Therefore in this case, the computer-generated results appear to be more accurate than the human-coded answer.

Criterion 5: Hematologic status (under <u>special</u> <u>procedures at admission</u>). Here the retrieval subroutine requires that tests involved in determining hematological status be recorded as being done at admission. Although the hematology sub-paragraph under LABORATORY DATA does contain the correct set of tests, no time is given. Therefore these answers are not accepted by the retrieval routine as meeting criterion 5.

Criterion 7: Blood culture (under <u>special proce-</u> <u>dures at admission</u>). This case is similar to the problem discussed for criterion 5. The sub-paragraph BACTERIOLOGY under LABORATORY DATA contains the sentence <u>3 blood cultures negative</u> with no indication of when the cultures were performed. The failure of the document to meet the time specifications in criteria 5 and 7 raises an interesting set of questions and possibilities:

a. Document deficient.

Are the documents themselves deficient for not stating the time of the lab test explicitly (in which case it is correct that the documents should "fail" the criteria)?

b. Implementation too literal.

Can we safely assume that the first set of lab tests were done at admission if not otherwise mentioned? In this case, the retrieval routine should be changed to look for the first set of laboratory tests instead of requiring that one set of tests be explicitly associated with the time of admission.

c. Criteria too stringent.

Are the criteria stated too stringently? Would it give more useful screening results if it were required that these tests be done sometime during the hospitalization, but not necessarily at admission?

Criterion 8: Reduction or disappearance of pain and swelling (under discharge status). In this case a negative answer was obtained by the computer because only reduction of pain was found. The actual text occurrences that produced these results are shown in another supplementary computer output (Figure 3). This output traces the progress of the retrieval subroutines through the text and prints a message each time a criterion (or a subpart of a criterion) has been found in the text. In particular the figure shows this trace over several paragraphs of the document from Fig. 1. From the third message we find that criterion 2: swelling of hands and/or feet is satisfied by the occurrence of her hands remained somewhat warm and swollen. The fourth message shows that criterion 1 is satisfied by however pain seemed to subside; it is satisfied because a statement that pain has subsided is taken as an implicit statement that pain existed. Note that the routine can identify the pain as being in the hands, because hand plural has been filled in from context in generating the data base (see Table 1, line 7). Finally this same sentence also satisfies part of criterion 8: disappearance/reduction of pain. But the retrieval routine fails to identify the sentence slight residual swelling of both hands as satisfying the criterion disappearance/reduction of swelling. This is the result of limitations in the retrieval routine (discussed below); the physician reviewer concluded from the text that both swelling and pain showed a reduction or disappearance, as required by criterion 8 of Fig. 2.

The operation of the retrieval routine for criterion 8 is outlined in Figure 4. To find an instance of reduction/disappearance of pain, the routine searches the data table for a row with an EXTREMITY-WORD in the BODY-PART column, a PAIN-WORD in the SIGN-SYMPTOM column, and an indication of reduction or disappearance which can be expressed in one of several ways:

A) there can be a word in NORMALCY, indicating a normal state or a change towards normal, as in swelling in hands cleared;

B) there can be a REDUCE-WORD in the QUANTifier column, e.g., less pain in hands;

C) there can be a REDUCE-WORD in the CHANGE column (which marks a change of state, e.g., reduction, but not the termination of that state), for example, pain decreased;

D) there can be a REDUCE-WORD in the END column (which marks the termination of a state, e.g., disappearance), as in pain subsided.

The fifth message in Fig. 3 shows that the sentence <u>pain seemed to subside</u> (line 7 in Table 1) meets the criterion <u>disappearance/reduction of</u> <u>pain: hand</u> is in BODY-PART, <u>pain</u> is in SIGN-SYMP-TOM, and <u>subside</u> is in END (alternative D in Fig. 4). However the retrieval routine for <u>reduction/</u> <u>disappearance of swelling</u> does not find any entries which satisfy its conditions. There are two sentences in the document that concern a reduction of swelling: <u>her hands remained somewhat warm and</u> <u>swollen</u> (Table 1, line 6) and <u>slight residual</u> <u>swelling of both hands</u> (Table 1, line 9). Neither of these is picked up by the retrieval routine.

FIGURE 3

PORTION OF COMPUTER-GENERATED MESSAGES

showing text sequences meeting the criteria of Fig. 2

			CRITERION	
			\sim	
(EL 2 - DIAGNOS) (FOUND FROM (* 1	IS - SWELLING OF EXPDS 3 1 5 (DOR	HANDS AND/OR FEET OR EQUIVALENT) SAL SURFACE AND PROXIMAL PALMAR SURFACE WAI	CS OF BOTH HANDS AF	E SWOLLEN,
(EL 1 - DIAGNOS) (FOUND FROM (*)	IS - PAIN IN HAN EXPDS 3 1 5 (DOR	DS AND/OR FEET OR EQUIVALENT) SAL SURFACE AND PROXIMAL PALMAR SURFACE WAN	CS OF BOTH HANDS AF M AND PAINFUL)))	E SWOLLEN,
(EL 2 - DIAGNOS) (FOUND FROM (* (IS - SWELLING OF COPDS 3 1 3 (HER	HANDS AND/OR FEET) HANDS REMAINED SOMEWHAT WARM AND SWOLI SUI	EN; HOWEVER PAIN S SSIDE)))	EEMED TO
(EL 1 - DIAGNOS) (FOUND FROM (* (IS - PAIN IN HAN COPDS 3 1 3 (HER	DS AND/OR FEET OR EQUIVALENT) . HANDS REMAINED SOMEWHAT WARM AND SWOLI SUI	EN; HOWEVER PAIN S SSIDE)))	EEMED TO
(EL 8 - REDUCTIO	ON/DISAPPEARANCE	OF PAIN)		
(FOUND FROM (* (COPDS 3 1 3 (HER	HANDS REMAINED SOMEWHAT WARM AND SWOLI	LEN; HOWEVER PAIN S SSIDE)))	EEMED TO
		N FOR ON COINC ORPEN		
(FOUND FROM (*)	RGE STATUS - PLA PLPDS 3 1 1 (NO	MEDICATIONS REQUIRED)))		
		FIGURE 4		
		RETRIEVAL ROUTINE		
۵.		ISADDEADANCE OF DEDUCTION OF DATH IN SUT		
	LOCATE AN ENTRY (POW) IN THE DATA TABLE THAT MEETS ALL OF (REMITIES,	
	1. BODY-PART HA	AS AN EXTREMITY-WORD:	RITERIA 1-3 BELOW.	
	2. SIGN-SYMPTON	HAS A PAIN-WORD:		
	3. ONE OF A,B,C	, OR D:		
	A. NORMALCY	IS FILLED:		
	B. QUANT	HAS A REDUCE-WORD;		
	C. CHANGE	HAS A REDUCE-WORD;		
	D. END	HAS A REDUCE-WORD;		
	4. NO NEGATION	IS ASSOCIATED WITH THE ENTRY (NEG COLUMN E	MPTY);	
	5. TIME ASSOCIA OR AT DISCHA	TED WITH THE ENTRY IS DURING CURRENT HOSPI RGE (TIME = <u>DISCH</u>).	TALIZATION (TIME = E	<u>IOSP</u>).
	WORD CLASSES:			
	EXTREMITY-WORD:	HAND, FOOT, EXTREMITY, FINGER, TOE, FORE	RM,	
	PAIN-WORD:	PAIN, PAINFUL, TENDER, TENDERNESS, SORE,	DISTRESS,	
	REDUCE-WORD:	REDUCE, REDUCTION, LESS, LESSEN, DISAPPEA	R, SUBSIDE,	
		DIMINISH, DECREASE,		
	NORMALCY	CONTAINS THE FOLLOWING WORDS: NORMAL, CLEAR, OK, FINE, GOOD, HEAL, IMPR	OVE, INTACT,	
в.	CRITERION 8.2. D	ISAPPEARANCE OR REDUCTION OF SWELLING IN E	XTREMITIES.	
	IDENTICAL TO PART	A EXCEPT THAT SWELLING-WORD IS SUBSTITUTE	D FOR PAIN-WORD IN 2	
	WORD CLASSES:			

SWELLING-WORD: SWELLING, SWELL, SWOLLEN.

The first sentence contains a genuine ambiguity: it is not clear whether somewhat modifies warm only or swollen as well in somewhat warm and swollen. In Table 1, it has been interpreted as modifying only warm. If it modifies both words, however, then it does indicate a reduction of swelling relative to the original mention, palmar surfaces are warm, swollen, and painful (HISTORY paragraph). However, this comparative reduction in swelling will not be retrieved by the routine as presently formulated. Similarly, the second sentence concerning swelling (slight residual swelling of both hands) is not retrieved because slight (like somewhat) implies a reduction only if compared to an original larger quantity. In addition, residual was not listed as a REDUCE-WORD. In order to capture these expressions of reduction of swelling, the program as outlined in Fig. 4 would have to be modified.

From this discussion of the discrepancies between the computer-generated and the hand-coded evaluations, we see that the 4 discrepancies can be summarized as follows:

1 case where the computer answer was more accurate than the hand-coded answer;

l case where the computer-generated answer was incorrect, because the retrieval routine failed to cover a complex way of expressing certain information;

2 cases where the discrepancy was due to the fact that the computer routine took a more strict interpretation of the criteria than the human coder, pointing to a need to clarify or change the criteria.

Discussion

Although there are discrepancies between the computer-generated evaluation and the hand-coded results, this experiment indicates that the computer-generated answers were, on the whole, comparable in reliability to the hand-coded results. By sufficient refinement of the retrieval routines, it may well be possible to produce computer evaluations that are more accurate and certainly more consistent than the hand-coded evaluations.

The existence of a computer program to assist in medical audit makes it possible to consider a systematic comparison of different sets of criteria used in health care evaluation. Once a data base of processed patient summaries has been prepared, it would, for example, be possible to test and evaluate sets of alternate criteria on this data base.

Alternate criteria can be programmed with relative ease by drawing on a library of subroutines. For example, the reduction or disappearance of any specified symptom can draw on the reduction/disappearance module outlined in step 3 of Fig. 4. There is also a time module that determines the time of an event relative to another event. Another module determines whether a condition is asserted, doubted, or denied, e.g., <u>patient has osteo-</u> myelitis/ <u>suspicion of osteomyelitis/ no evidence</u> <u>of osteomyelitis</u>. Often all that is required is to provide synonym lists for the desired entries; these are plugged into a set framework drawing on the library subroutines, to produce the desired routine.

The computer application in health care evaluation described here has not as yet been engineered for routine use in a hospital setting, although this is one of our present goals. It is part of a larger research effort of the N.Y.U. Linguistic String Project, aimed at improving access to information in natural language documents, in particu-lar, narrative patient records.^{5,6} For example, the Linguistic String Project is presently engaged in an experiment on automatic coding of symptoms of cancer of the head and neck. The data is recorded in free-text form as a part of a data base on head/neck cancer maintained by Roswell Park Memorial Institute. In this case, the program produces a code for each recorded symptom and its location. This program is currently being tested and compared to the hand-coded results.

Implementation

The program is implemented on a Control Data 6600 and requires approximately 75,000 words of memory. The formatting component (which creates the data table) is implemented in FORTRAN, while the retrieval program is written in LISP 1.5. The processing times for the document shown in Fig. 1 were approximately 9 minutes for the creation of the data base, 5 minutes for normalization (filling in implicit contextual information and time information) and 2 minutes for retrieval (evaluation of the health care criteria).

Acknowledgments

This research was supported in part by NIH Grant LM02616 from the National Library of Medicine.

The LISP programming for the normalization and retrieval was done by Guy Story. Elaine Marsh assisted in processing the documents for the creation of the data base.

References

- McDermott, W., Evaluating the Physician and his Technology. <u>Daedalus, The Journal of the American Academy of Arts and Sciences</u>, Winter 1977, 135-157.
- Brook, R.H., and F.A. Appel, Quality-of-Care Assessment: Choosing a Method for Peer Review. <u>New England Journal of Medicine</u> 288:25, 1323-1329 (1973).
- Hirschman, L., Grishman, R., and N. Sager, From Text to Structured Information: Automatic Processing of Medical Reports. <u>AFIPS Conference</u> <u>Proceedings</u> <u>45</u>, 267-275, AFIPS Press, Montvale, N.J., 1976.
- Sager, N. and L. Hirschman, Information Structures in the Language of Science; Theory and Implementation. <u>String Program Reports (S.P.R.)</u> No. 12, Linguistic String Project, New York University, 1978.
- Sager, N., Natural Language Information Formatting: the Automatic Conversion of Texts to a Structured Data Base. <u>Advances in Computers 17</u>, 89-162, Academic Press, N.Y., 1978.

6. Sager, N., Hirschman, L., and M. Lyman, Computerized Language Processing for Multiple Use of Narrative Discharge Summaries. <u>Proceedings of the Second Annual Symposium on Computer Application in Medical Care</u>, 330-343, IEEE, New York, 1978.