Natural Language Processing of Asthma Discharge Summaries for the Monitoring of Patient Care

Naomi Sager [1], Margaret Lyman [2], Leo J. Tick [2], Ngô Thanh Nhàn [1], Christine E. Bucknall [3]

[1] Courant Institute of Mathematical Sciences, New York University, New York, NY 10012
[2] New York University Medical Center, New York, NY 10016
[3] Area Clinical Audit Office, Dept. of Public Health, Greater Glasgow Health Board, Glasgow, Scotland

ABSTRACT

A technique for monitoring healthcare via the processing of routinely collected narrative documentation is presented. A checklist of important details of asthma management in use in the Glasgow Royal Infirmary (GRI) was translated into SQL queries and applied to a database of 59 GRI discharge summaries analyzed by the New York University Linguistic String Project medical language processor. Tables of retrieved information obtained for each query were compared with the text of the original documents by physician reviewers. Categories (unit = document) were: (1) information present, retrieved correctly; (2) information not present; (3) information present, retrieved with minor or major error; (4) information present, retrieved with minor or major omissions. Category 2 (physician "documentation score") could be used to prioritize manual review and guide feedback to physicians to improve documentation. The semantic structuring and relative completeness of retrieved data suggest their potential use as input to further quality assurance procedures.

INTRODUCTION

The current emphasis on cost containment and quality assurance has made the assessment of the healthcare process an element in patient care, almost as important as the care itself. Urgently needed are computer-based methods for obtaining the patient data as they relate to established healthcare guidelines and for verification of such guidelines in outcomes research. This is especially timely in light of the movement toward adopting a computer-based patient record as a national standard [1]. Current audit practices require a trained individual to read the medical records of a subset of patients in order to extract the necessary data. With the development of natural language technology, tools are available for performing some portion of this task by processing routinely collected narrative accounts of patient care, e.g. radiology reports, admission notes, discharge summaries.

Natural language processing (NLP) systems have been slow to develop. The inherent complexity of the task, along with such specific problems of processing as the resolution of syntactic ambiguity and the balancing of semantic adequacy vs. system robustness, has impeded development toward applications [2]. General purpose NLP systems must be tailored for the domain of application; dictionaries and thesauri must be created. Fortunately, within medical informatics considerable attention is being paid to these tasks and systems are emerging [3, 4, 5, 6].

The Linguistic String Project of New York University has developed a medical language processor (LSP-MLP) that analyzes free-text patient documents in such manner as to make the information contained in the documents accessible by queries to a database [7]. The present paper reports on the use of the system to extract data from 59 hospital discharge summaries of asthma patients at the Glasgow Royal Infirmary (GRI), according to the requirements of a checklist of important details of asthma management developed there. A reasonable hope would be that, with the aid of the computer processing of patient documents, patient care audit would become a routine, instead of as now, an episodic activity.

METHODS

A study of in-hospital management of asthma in relation to outcome as measured by post-discharge interview was conducted at the Glasgow Royal Infirmary (GRI) [8, 9]. A checklist of important details of asthma management was developed, which has been utilized in this investigation [10]. The checklist contains 6 major categories of information, subdivided to form 13 checklist items in all (listed at the left in

Table 1, below). These were formulated as SQL queries to operate on the database which was derived from 59 GRI hospital discharge summaries, using the LSP-MLP system.

The LSP program is composed of five modules that operate in sequence on each successive sentence of a given document to: (1) determine syntactic structure (parsing); (2) resolve ambiguity and semantically label parse tree substructures; (3) regularize parse tree substructures into elementary assertions and connectives; (4) convert connective structure into Polish notation; (5) link semantically-labelled nodes of the final sentence tree with corresponding nodes of the medical representational structure. Finally, the results of processing a set of documents are mapped into a database management system, currently a relational DBMS [11, 12].

To ensure that medical language processing when applied to patient documents produces reliable patient data, the LSP system contains a "NIMF" quality control procedure, which is used with each run:

(1) A database field holds the results of automatic quality assessment of the row to be loaded. It contains

empty	if the row passes all tests;
N	if there is a potential Negative problem;
I	if the row is semantically Ill-formed (wrong type word in field);
M	if there is a potential Modal problem (Modal=uncertainty);
F	if there was a processing Failure (the whole sentence is in the TEXTPLUS field).

(2) Rows with a non-empty quality assessment field are not loaded; the sentences with such problems are rerun using a modified parsing procedure that recovers wellformed rows from analyzable parts of sentences. All the original text material is in the database, either analyzed or in TEXTPLUS.

RESULTS

This section shows the results of retrievals for the checklist items deemed important in asthma management. Tables of retrieved information obtained for each query were compared with the text of the original documents by physician reviewers (Drs. Bucknall and Lyman).

Figure 1 shows for 4 patients the sections of the tables generated by SQL queries for checklist items 2 (Time Since Asthma Well Controlled) and 3c (Abnormal Findings on Admission) along with the original contributing sentences. Item 2 was interpreted to mean that specific time information was present in the document with regard to at least one of the presenting symptoms (e.g. 056B.1.01). Item 3c was evaluated with respect to the report of the admission physical examination. Since several sentences often were involved, all findings from all sentences were considered. Error was counted as minor if a misplaced word did not destroy the main meaning; a major error could be retrieving as an admission finding one that was not at admission. A "miss" was considered minor if most of the findings were retrieved (as in Fig. 1, where 056C.1.03 misses RHONCHI) and major otherwise.

Negative findings are identified in the processing and eliminated in the retrieval. Thus, in Fig. 1, sentence 047B.1.02 there are no rows corresponding to "though she had no problems with cough or sputum production"; cf. also 002C.1.05, 046C.1.01, 056C.1.03.

Table 1 summarizes the results, using the document as the unit of measurement. For example, for checklist item 1 (Therapy before Admission), 15 of the 59 discharge summaries did not contain this information (Column 2). The LSP-MLP system (including SQL queries) correctly retrieved all such information from 31 documents (Column 1), retrieved the desired information from 6 documents but with some errors in the report of the information (Column 3), and missed some or all of the desired information in 7 documents (Column 4). The system did not retrieve anything from any of the documents that did not contain the desired information.

To evaluate the effectiveness of the system we used columns 1, 3, and 4 since the values in column 2 reflect the number of documents that contained no information for the system to retrieve. New measures will have to be developed for situations like ours where the dimensions/variables include whether a document containing information was located (appeared in retrieval results), whether the information was reported correctly in the retrieval results and whether the reported information was complete. Major vs. minor departures from total correctness must be defined for each application.

As a start, we define Information Precision (I-P) for each query as the number of documents for which the desired information was retrieved divided by the total

Table 1: Checklist of Important Details of Asthma Management
Summary of Computer-Retrieved Information (59 Discharge Summaries)

	NUMBER OF DOCUMENTS			
	Information present; Retrieved correctly	**Information not present**	**Information present; Retrieved with some error ***	**Information present; Missed in part or whole ****
1. Therapy before Admission	31	15	(5,1)	(5,2)
2. Time since Asthma Well Controlled	33	20	(0,0)	(0,6)
3. Admission				
a. Pulse	42	16	(0,1)	(0,0)
b. Peak Flow Rate	17	27	(8,4)	(1,2)
c. Abnormal Findings	41	3	(11,0)	(3,1)
d. Chest X-rays	26	19	(3,0)	(4,7)
e. Blood Gases	29	17	(3,0)	(8,2)
4. Treatment Given Including Oxygen	38	0	(16,2)	(1,2)
5. Discharge:				
a. Peak Flow	15	42	(0,0)	(1,1)
b. Repeat Blood Gases, if done	6	52	(0,0)	(0,1)
6. Arrangements at Discharge				
a. Prednisolone	32	15	(3,0)	(4,5)
b. Long-term Therapy	43	7	(3,0)	(2,4)
c. Review Arrangement	33	6	(9,0)	(5,6)

* In this column (n_1,n_2): n_1 = minor error, n_2 = major error.
** In this column (n_1,n_2): n_1 = minor miss, n_2 = major miss.

number of documents for which any information was retrieved:

$$\text{I-P} = \frac{\text{value in column 1}}{\text{sum of values in columns 1 and 3}}$$

Similarly, we define Information Recall (I-R) for each query as the number of documents for which the desired information was retrieved divided by the total number of documents that contained such information:

$$\text{I-R} = \frac{\text{value in column 1}}{\text{sum of values in columns 1 and 4}}$$

Table 2 shows I-P and I-R for the 13 queries. Average I-P was 85.9% and average I-R was 84.2%. When I-P and I-R were calculated using the counts of major errors and major omissions only, the respective scores were 98.3% and 91.0% (average major error 1.7%, average major miss 9.0%).

Table 2: Some Performance Measures of Medical Language Processing and SQL Retrieval

Checklist item	Information Precision (%)	Information Recall (%)
1	83.8	81.6
2	100.0	84.6
3a	97.7	100.0
3b	58.6	85.0
3c	78.8	91.1
3d	89.7	70.3
3e	90.6	74.4
4	67.9	92.7
5a	100.0	88.2
5b	100.0	85.7
6a	88.9	78.0
6b	93.5	87.7
6c	67.3	75.0
Average	**85.9**	**84.2**
Average major errors/ misses only	**98.3**	**91.0**

DISCUSSION AND CONCLUSIONS

Table 1, column 2 provides data on deficits in the quality of documentation of patient care with respect to stated criteria. If one considers the items not mentioned to be important in asthma management, one can see immediately where emphasis on teaching could be placed, or feedback to the reporting physicians introduced. In the real world of audit, knowing which discharge summaries had significant deficits in documentation would allow the human reviewer to be selective with regard to which records need manual review.

In terms of the potential use of language processing as an aid in the audit task, the language processing tool demonstrated here has several significant features:

1. All documents are treated consistently with regard to a given criterion.
2. Only significant information is retrieved.
3. Major errors in retrieval results are minimal (average 1.7% in this experiment).
4. Major omissions in the retrieval results are relatively small in number (average 9.0% in this experiment).
5. The semantic structuring and relative completeness of retrieved data suggest their potential use as input to further quality assurance procedures.

The LSP-MLP was designed as a multiple purpose device. The database created from the documents can be used for purposes other than audit, for example, to determine from a set of analyzed documents for a given diagnosis those cases which are most similar in regard to physical and history findings [13]. More generally, it has been our purpose to provide a means whereby the detailed clinical context expressed in narrative reports can be incorporated into the computer-based patient record, and thereby not lost to further decision procedures and retrospective review.

REFERENCE

[1] Dick, R.S. and Steen, E.B. (eds), *The Computer-based Patient Record*, National Academy Press, Wash. D.C., 1991.

[2] Grosz, B.J., Jones, K.S., and Webber, B.L. (eds), *Readings in Natural Language Processing*, Morgan Kaufman, Los Altos, CA, 1986.

[3] Humphreys, B.L. and Lindberg, D.A.B., The Unified Medical Language System Project: A Distributed Experiment in Improving Access to Biomedical Information, *MEDINFO 92*, North-Holland, Amsterdam, 1992.

[4] Gabrielli, E R, Computer assisted assessment of patient care in the hospital, *Journal of Medical Systems* 12:135, 1988.

[5] Vries, J.K., Marshalek, B., D'Arbarno, J.C., Yount, R.J. and Dunner, L., An automated indexing system utilizing semantic net expansion, *Computers and Biomedical Reseach* 25:153-167, 1992.

[6] Lin, R., Lenert, L., Middleton, B. and Shiffman, S., A free-text processing system to capture physical findings: Canonical Phrase Identification System (CAPIS). *SCAMC15*, McGraw Hill. pp. 843-847, 1991.

[7] Sager, N., Friedman, C., Lyman, M.S., and members of the Linguistic String Project. *Medical Language Processing: Computer Management of Narrative Data*. Addison-Wesley, 1987.

[8] Bucknall, C.E., Robertson, C., Moran, F. and Stevenson R.D. Differences in hospital asthma management. *Lancet* 1:748-750, 1988.

[9] Bucknall, C.E. et al. Management of asthma in hospital: a prospective audit. *Brit Med J* 296:1637-1639, 1988.

[10] Lyman, M., Sager, N., Nhàn, N.T., Tick, L.J., Borst, F. and Scherrer, J-R. The Application of Natural-Language Processing to Healthcare Quality Assessment, *Medical Decision Making* 11, #4 Suppl., pp. S65-S68, 1991.

[11] Nhàn, N.T., Sager, N., Lyman, M., Tick, L.J., Borst, F. and Su, Y. A Medical Language Processor for Two Indo-European Languages. *SCAMC13*, L.C. Kingsland, ed., Washington, D.C., pp. 554-558, 1989.

[12] Sager, N., Lyman, M.S., Tick, L.J., Nhàn, N.T., Borst, F. and Scherrer, J-R., Clinical Knowledge Bases From Natural Language Patient Documents, *MEDINFO 92*, North-Holland, Amsterdam, 1992.

[13] Borst, F., Lyman, M.S., Nhàn, N.T., Tick, L.J., Sager, N. and Scherrer, J-R. TEXTINFO: A Tool For Automatic Determination of Patient Clinical Profiles Using Text Analysis, *SCAMC15*, McGraw Hill. pp. 63-67, 1991.