

Research Paper ■

An Experiment Comparing Lexical and Statistical Methods for Extracting MeSH Terms from Clinical Free Text

GREGORY F. COOPER, MD, PHD, RANDOLPH A. MILLER, MD

Abstract **Objective:** A primary goal of the University of Pittsburgh's 1990–94 UMLS-sponsored effort was to develop and evaluate PostDoc (a lexical indexing system) and Pindex (a statistical indexing system) comparatively, and then in combination as a hybrid system. Each system takes as input a portion of the free text from a narrative part of a patient's electronic medical record and returns a list of suggested MeSH terms to use in formulating a Medline search that includes concepts in the text. This paper describes the systems and reports an evaluation. The intent is for this evaluation to serve as a step toward the eventual realization of systems that assist healthcare personnel in using the electronic medical record to construct patient-specific searches of Medline.

Design: The authors tested the performances of PostDoc, Pindex, and a hybrid system, using text taken from randomly selected clinical records, which were stratified to include six radiology reports, six pathology reports, and six discharge summaries. They identified concepts in the clinical records that might conceivably be used in performing a patient-specific Medline search. Each system was given the free text of each record as an input. The extent to which a system-derived list of MeSH terms captured the relevant concepts in these documents was determined based on blinded assessments by the authors.

Results: PostDoc output a mean of approximately 19 MeSH terms per report, which included about 40% of the relevant report concepts. Pindex output a mean of approximately 57 terms per report and captured about 45% of the relevant report concepts. A hybrid system captured approximately 66% of the relevant concepts and output about 71 terms per report.

Conclusion: The outputs of PostDoc and Pindex are complementary in capturing MeSH terms from clinical free text. The results suggest possible approaches to reduce the number of terms output while maintaining the percentage of terms captured, including the use of UMLS semantic types to constrain the output list to contain only clinically relevant MeSH terms.

■ JAMIA. 1998;5:62–75.

Affiliations of the authors: Center for Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA (GFC); Division of Biomedical Informatics, Vanderbilt University, Nashville, TN (RAM).

Supported by UMLS contract N01-LM-3535 from the National Library of Medicine.

Correspondence and reprints: Gregory F. Cooper, MD, PhD, University of Pittsburgh, Center for Biomedical Informatics, Suite 8084 Forbes Tower, 200 Lothrop Street, Pittsburgh, PA 15213-2582. e-mail: <gfc@cbmi.upmc.edu>.

Received for publication: 8/6/97; accepted for publication: 9/17/97.

The body of medical literature continues to expand at a rapid pace, and it is increasingly difficult for clinicians to keep track of all the literature that may have a significant influence on the care of their patients.¹ At the same time, an increasing amount of patient clinical information is being stored in computers. A key goal of the National Library of Medicine's Unified Medical Language System (UMLS) project is to assist healthcare practitioners, researchers, and students to link the patient medical record to the medical literature.² As part of the UMLS Project from 1990–1994, we developed initial versions of two computer pro-

grams, called Pindex and PostDoc.³⁻⁵ Each program takes as input a portion of the free text from a narrative part of a patient's electronic medical record, and it returns a list of suggested MeSH terms to use in formulating a Medline search that includes concepts in the text. The Pindex program uses a probability-based statistical method to perform this mapping, whereas PostDoc uses a simple lexical matching method. We believe that a system that suggests a list of potentially relevant MeSH terms can help clinicians improve the number of relevant Medline articles retrieved. In this paper we do not experimentally address whether such an improvement occurs in a clinical setting; rather, we report the results of a formative study that bears on the issue in an important way. In particular, this paper reports the results of an initial evaluation of the ability of the two systems to map the concepts in clinical reports to MeSH. The intent is for this evaluation to serve as one step toward the eventual realization of systems that assist healthcare personnel in using the electronic medical record to help construct patient-specific searches of Medline.

Background

The ability to map accurately from clinical free text to a structured vocabulary has important uses in the indexing and retrieval of medical literature. Researchers have begun to explore this mapping process using the UMLS Metathesaurus. A substantial amount of research on concept recognition within texts, and specifically within clinical texts, has been done⁶⁻¹⁶; additional relevant references with abstracts appear in a UMLS bibliography, which was compiled by Selden and Humphreys.¹⁷ Examples of relevant prior research include work on indexing medical images,¹² indexing clinical articles based on their abstracts and retrieving them based on a free-text phrase,⁹ mapping ICD-9 terms to MeSH terms,¹⁰ and lexical rules for linking clinical reports to the medical literature.⁸ The current paper empirically investigates two methods (and their combination) for mapping from the text in an entire clinical report to the MeSH terms that represent it. In the remainder of this section, we describe the two methods and their implementation. We also suggest how the methods might be used in a system that assists users in searching Medline for articles that are specific to the care of individual patients.

The PostDoc Lexical Recognition Algorithm

The PostDoc algorithm was originally developed in 1991-92.⁴ The PostDoc approach was based on the "keep it simple" principle. The justification for a "simple" approach was recognition that the goal was *not*

to "understand" fully the content of arbitrary medical text. The goal was to extract, as best as possible, references to terms in a specified target lexicon (UMLS Metathesaurus/MeSH) as they appeared in an arbitrary medical text (so long as the mappings could be justified clinically). Given that the task was simplified to mapping arbitrary text to a target lexicon, the algorithm was developed on the assumption that the medically meaningful content in free-text clinical records would be contained within noun phrases, because the target vocabulary for matching was a controlled vocabulary of noun phrases (the UMLS Metathesaurus lexicon, and its subset, MeSH). Another PostDoc assumption was that all the important medical words worth recognizing in free-text noun phrases should be related to (i.e., derived from) the words in the target vocabularies (including synonyms and lexical variants). Recognizing a word that does not appear anywhere in a target lexicon (including synonyms) is less likely to be useful, and knowing what to do with such words would require substantial manual effort. It was our experience at the time that the majority of words in noun phrases found in medical charts are "medical words," in that they are words that participate in terms of the UMLS Metathesaurus. The set of target words was constructed by taking the root (canonical) form of each unique word that participated in any UMLS Metathesaurus (including MeSH) term name, or in the term name's lexical and linguistic variants (including synonyms). For historical purposes, notice that the early (1990-92) versions of the UMLS Metathesaurus did not contain root-form versions of words participating in UMLS terms, even though these roots are provided in the current releases of the Metathesaurus.

Notice that the PostDoc algorithm could handle synonymy in most, but not all cases. Consider a primary Metathesaurus term containing the word "hepatic." If the Metathesaurus lexical variants for that primary term included substitution of the word "liver" for "hepatic" in some of the lexical variants, PostDoc could map chart text phrases containing "liver" to the primary Metathesaurus concept (if justified) by first mapping to the Metathesaurus lexical variant form, and then mapping from a variant Metathesaurus term to the primary (preferred) Metathesaurus concept term. However, consider a second example. If none of the Metathesaurus lexical variants of a preferred Metathesaurus term containing the word "hepatic" contained the word "liver," PostDoc could not map a chart phrase containing "liver" to the preferred term containing the word "hepatic," since PostDoc did not perform word-level synonymy checking. Generally, not all instances of "liver" occurring in arbitrary text would map to "hepatic," and the lexical variants in-

cluded in the Metathesaurus should be a good way to determine when such mappings make sense. Thus, the ability of PostDoc to carry out synonymy mapping was dependent on the quality of the lexical variant terms included in the Metathesaurus.

To create PostDoc data structures, we processed the 1991–92 Meta-1.1 CD-ROM data files. We extracted all unique words from terms in the UMLS “MRMC” (main concepts) file. We next created inverted indices (Key Word in Context, or KWIC lists) to tie each word back to the UMLS main-concept terms in which they appeared. It was important to store the unique words and the inverted word indices based on “root” forms of each word. While sophisticated methods exist to create a linguistic “stem” for a given word,^{18,19} the simple approach taken in PostDoc was to convert all words to a pseudo-singular form. We accomplished this by writing a program that recognized (and could generate) potential English, Latin, or Greek singular, possessive, or plural forms of words. Only the singular form of a given word was stored as its “root string,” and the inverted indices of variant forms of the root string were all merged into a single index per word. The algorithm used to map words to their root form was also required as part of the run-time PostDoc analysis routines, since the words in a free-text patient record may appear arbitrarily in singular, plural, or possessive form.

PostDoc employed a simple, “sliding frame” method for matching phrases in the source document to those in the target lexicon. Three passes were made by the program. The first pass took all words in the source document sequentially, and outputted only recognized words in their “root” forms. In the second pass (described below), the words output by the first pass were grouped into potential matches for target lexicon terms. In the third pass (described below), the output of the second pass was evaluated and finalized.

In analyzing the words in a free-text clinical document, the PostDoc algorithm matched any “word” (group of characters set off by “whitespace” delimiters) in the source document with entries in the inverted word indices that were either identical in length and content to the source word’s root form or identical up to the length of the source word but with the length of the index word being longer (containing additional characters). For this reason, it was decided to ignore source document words shorter than a given length, which was set to be four characters. During PostDoc development, it was empirically observed that shorter words matched too many root words in the index nonspecifically, and resulted in too many false-positive matches. We also made an arbitrary de-

cision, again based on empirical observations, to treat certain common English words as “stop words” that would be dropped from the recognition process as if they had never appeared in the source text. The stop words list included: after, and, brief, course, date, day, down, felt, focus, in, mild, minute, minutes, of, other, second, seconds, than, the, these, this, time, times, week, when, will, with, year. Because the algorithm used inverted word indices, word-order variations between source phrases and target phrases were permitted. When stop words were implemented, PostDoc allowed terms such as “carcinoma of the prostate” and “prostate carcinoma” to match as being identical terms.

In its second phase of processing, PostDoc used a serial sliding-frame method to accumulate potential matches for source-document phrases (sequential combinations of words). The pseudo-code version of the PostDoc phase 2 algorithm is given in Appendix A.

In PostDoc’s third and final phase, a set of heuristic rules²⁰ was applied to the output from phase 2 to see whether an appropriate match existed between the words in the chart and one (or a few) Meta-1.1 term(s). In effect, PostDoc had to determine whether it was reasonable to match a series of “recognized” words from the chart to the set of Meta-1.1 terms corresponding to those words’ non-null intersected KWIC lists. The first heuristic used in PostDoc phase 3 counted the number of chart words that actually appeared in each candidate Meta-1.1 term “matched” (i.e., the output of phase 2). If less than 51% of the words in the candidate Meta-1.1 term appeared in the chart, the term was rejected. In this manner, the isolated word “carcinoma” in a chart could not match all Meta-1.1 terms containing the word “carcinoma.” Similarly, even though the solitary word “diabetes” in an arbitrary medical text usually refers to “diabetes mellitus,” it is not medically justified to make this mapping automatically, since diabetes insipidus (of pituitary origin) and nephrogenic diabetes insipidus also include the word “diabetes.” In addition, only the candidate Meta-1.1 term with the highest percentage of matched words from each chart phrase was retained. For example, if a phrase found in a chart was “insulin-dependent diabetes,” then 75% of the words in the Meta-1.1 term “Diabetes Mellitus, Insulin Dependent” were matched, but only 60% of the words in the term “Diabetes Mellitus, Non Insulin Dependent” were matched, so PostDoc dropped the latter term. Finally, a heuristic was employed to further limit “nonspecific” matches. It was determined to set a cutoff (at five) for the maximum number of candidate matches that PostDoc would retain for a phrase

in the chart (i.e., there must be five or fewer Meta-1.1 terms matching each phrase output by the PostDoc phase 2 algorithm given as pseudo-code in Appendix A).

Additional aspects of the PostDoc algorithm for identifying co-occurring terms were previously described.⁴

The Pindex Statistical Indexing Algorithm

In this section, we describe the methods used in Pindex to construct the associations between phrases and MeSH terms, and the techniques used to apply these associations to stochastically derive MeSH terms from the phrases in a section of clinical text. The basic method, which was developed initially in 1991 as part of the UMLS project, was subsequently extended as part of that project.

The Construction of Associations in Pindex

Pindex takes as input a string or file of free text, and it returns MeSH terms (MTs) that co-occur most frequently given the phrases in the text. The MTs used in the current version of Pindex are MeSH Main Headings. We constructed the associations between phrases and MTs by associating the phrases in Medline titles and abstracts with the MTs that human indexers at the National Library of Medicine (NLM) have attached to the respective Medline articles. Synonyms were not entered separately into the version of Pindex described here, but they often are represented by multiple free-text phrases that are strongly statistically associated with a given MT.

We used Medline articles from the first six months of 1990 to construct associations between MTs and the free-text phrases of titles and abstracts. There was a total of 193,160 articles (all containing titles), of which 133,876 (69%) also contained abstracts. We will use the term *article text* to refer to the available online text of a given article, which consists either of a title only, or of a title plus abstract.

Training occurred as follows. (See Appendix B.1 for a simple example that illustrates how Pindex was trained using the steps that follow.) For a given Medline article, we partitioned the article text into a set of subtexts by applying stop words, which served as barrier words.* We then transformed each subtext as follows.

- We removed all apostrophes (e.g., *Graves' Disease* became *Graves Disease*).

- We converted all letters to lower case (e.g., *Graves Disease* became *graves disease*).
- All non-alphabetic characters were replaced by blanks.
- Extra blank spaces were removed.

Now that we had the subtext in a standard format, we broke it into phrases. In particular, we determined every possible singleton, doublet, triplet, and quadruplet of words in the subtext. For example, a triplet phrase consists of three sequential words that occurred in the subtext. Thus, word order was important in constructing phrases. We did not consider phrases longer than four words, because phrases of one to four words seemed adequate to capture most associations. In the future, it may be worth investigating this assumption by using longer phrases.

We took the union of all the phrases in all the subtexts of a given article. This union represents the set of phrases P for the article. Notice that by using the union of phrases, a given phrase will appear only once in set P . Over all articles, we tallied the number of articles (P sets) in which a given phrase appeared. We did this for all phrases found in all articles. We permanently stored only phrases that occurred in at least two articles, because a single occurrence of a phrase is not enough to establish those MTs with which the phrase has a statistically meaningful association.

Each Medline article has a set of MTs that a human indexer at NLM has assigned to it. We did not consider most MTs that are check tags or similar terms, because this is a short list of terms that occur commonly, and thus, users can select such search terms directly from the list.† Let M denote the set of remaining MTs for an article. We took the cross product of the elements in P and M to construct a set PM . Thus, PM contained every possible combination of a phrase in P with a MT in M . Over all articles, we tallied the number of times a given phrase-MT pair occurred.

We used the frequency of a phrase X (let m denote this frequency) and the frequency of co-occurrence of a phrase-MT pair (let k denote this frequency) to calculate the conditional frequency of the MT being assigned to an article as an indexing term given that the phrase appears in the title or abstract of the article. Let Y denote the MT. If we interpret this conditional frequency as being a conditional probability, then we can compute the probability of Y given X as follows:

*A list of the 81 stop words that were used can be obtained from author G. F. Cooper.

†A list of the 28 MTs that were excluded can be obtained from author G. F. Cooper.

$P(Y|X) = k/m$. We retained in our database only those phrase-MT pairs for which $P(Y|X) \geq 0.1$. We used this filter because it significantly reduced the size of the database. We believe that using the filter has a negligible effect on indexing performance, because by informal inspection MTs with probabilities less than 0.1 are usually not useful indexing terms.

Table 1 shows the numbers of phrases and phrase-MT pairs that were captured from the six months of training data. More than 500,000 phrases and 2.7 million phrase-MT associations were captured. These

Table 1 ■

Numbers of Phrase and Phrase-MeSH Term (MT) Pairs Captured

	Total Phrases	Total Phrase-MT Pairs	Average MTs per Phrase
1-word phrases	61,355	408,384	6.7
2-word phrases	335,921	1,687,107	5.0
3-word phrases	116,410	514,435	4.4
4-word phrases	33,194	138,436	4.2
All phrase lengths	546,880	2,748,362	5.0

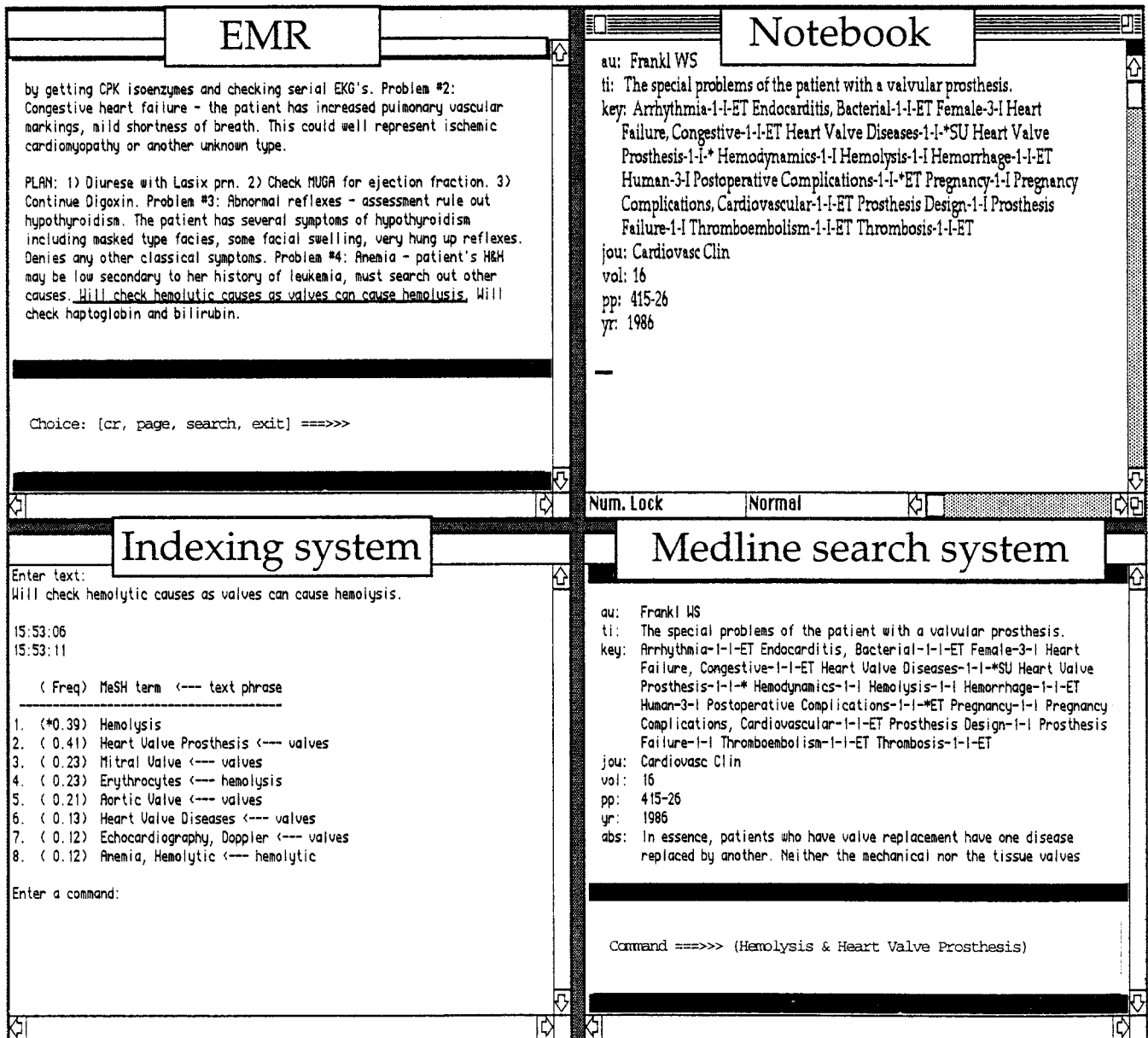


Figure 1 A prototype version of Chartline that uses Pindex as an indexing system.

data initially were placed in a text file, which we denote as file T .

The phrases and phrase–MT pairs in T were used to create a hash table H that uses a phrase as a key. Given a phrase, the computer can quickly determine whether the phrase is in the hash table, and if so, then it can efficiently access the MTs associated with that phrase (along with frequency of occurrence of the phrase and the frequency of co-occurrence of the associated MTs). The hash table resides on disk and occupies about 45 MB of disk space. Approximately 250 hours of dedicated time on an IBM RS/6000 workstation were used to construct the hash table from the six months of Medline articles; a substantial portion of this time was devoted to disk IO, since the hash table was too big to fit into the 16 MB of RAM available at the time.

The Application of Associations

The Pindex program takes as input a free-text phrase from the keyboard or from a file. The program then uses the methods described in the previous section to generate phrases from the text. For each phrase X , the MTs associated with that phrase are retrieved from hash table H . For each such associated MT Y , the program computes $P(Y|X)$. Over all the phrases generated from the text, the phrase X' that maximizes $P(Y|X')$ is noted, and this is output as the posterior probability of Y . All the MTs associated with the phrases in the article are sorted in descending order of probability. The MTs with a probability at least as great as a user-specified threshold t are displayed; in the experiments described in this article $t = 0.20$. Appendix B.2 demonstrates the application of Pindex to a portion of a medical chart.

The Chartline Concept

We use the term “Chartline” to designate a computer-based approach in which textual information in patient charts is mapped to clinical (Metathesaurus/MeSH) terms, and subsequently to the medical literature. We described a PostDoc-specific version of Chartline in 1992.⁴ A prototype interface to a version of Chartline that uses the PostDoc system has been developed, and it is described by Schwartz et al.²¹ In this paper, we suggest the future possibility of a version of Chartline that uses both the PostDoc and Pindex algorithms.

Figure 1 shows a working prototype of Chartline (using the Pindex engine), developed during the UMLS project in 1992–93. This prototype was not intended for clinical use per se and is shown here to convey the basic idea behind the Chartline approach.

Figure 1 shows the four windows on a hypothetical clinician user’s monitor. The top left window contains a patient record that a clinician might access using an electronic medical record system, such as the MARS system at the University of Pittsburgh,²² which is shown in the figure. Such a user would highlight (underline) text in a window that contains concepts that are of interest in performing a Medline search; the top left window shows such underlined text. The bottom left window shows search terms that were created by an indexing system, when applied to the underlined text. In particular, the output shown is MeSH terms that were produced by the Pindex system. An arrow indicates the text phrase that invoked a given MeSH term in the list.‡ The frequency of the term given the phrase also is shown. Suppose the user decides to use the MeSH terms *Hemolysis* and *Heart Valve Prosthesis* in performing a Medline search. The user selects these terms and applies them to construct the search expression that is shown in the bottom right window (lower part). The upper portion of the lower right window displays part of the output of a Medline search, which was produced by the MARS Medline search engine in this version of Chartline; other search systems, such as Grateful Med, could be substituted. The clinician has cut and pasted some of the Medline information about this article into a personal electronic notebook, which is shown in the top right window.

We believe the usefulness of the Chartline system will depend significantly on the quality of the MeSH terms returned by the system, when given free-text excerpts of clinical charts. To investigate this performance issue, we conducted a formative laboratory evaluation of the ability of Pindex and PostDoc to find the MeSH keywords that represent the significant clinical concepts in selected narrative sections of patient reports on MARS. In addition, we evaluated a hybrid version of PostDoc and Pindex that produces as its output the union of the output produced by PostDoc and Pindex.

Experimental Methods

For purposes of the study, the authors selected three representative, clinically common types of patient reports available through the MARS system: radiology reports, surgical pathology reports, and hospital dis-

‡The term *hemolysis* does not have an arrow pointing to it in Figure 1, because the identical text phrase *hemolysis* caused the term to be displayed there. The frequency of association between the phrase and the term is 0.39. An asterisk before the 0.39 indicates that the invoking phrase and the term are identical.

Sample Radiology Report

Report Type Radiology Report
 Diagnosis US/THYROID, (R) THYROID CYST*ENTC*XXXXX* 7/21
 Exam Type CUS635 THYROID/NECK

ULTRASOUND OF THE THYROID XXXXX 9:27AM

HISTORY: RIGHT THYROID CYST

Realtime ultrasonography was utilized in the evaluation of the thyroid gland. Transverse and sagittal images of the right lobe of the thyroid reveal two cyst-like structures with internal echos, one 8.5mm in diameter, the other 5.5mm in diameter. The left lobe of the thyroid contained a solid focal nodule measures 4.5mm in diameter immediately adjacent to the left carotid artery. The cystic structures on the right most likely represent an area of degenerating nodular goiter however the etiology of the focal nodule in the left lobe of the thyroid is more uncertain. If clinically warranted a Nuclear Medicine scan of the thyroid could be performed or alternatively a follow up study is 3 months time could be obtained.

1. TWO SMALL CYSTIC STRUCTURES IN THE RIGHT LOBE OF THE THYROID POSSIBLY REPRESENTING DEGENERATING NODULAR GOITER OR HEMORRHAGE INTO A CYST AFTER ASPIRATION.
2. FOCAL SOFT TISSUE NODULE IN THE LEFT LOBE OF THE THYROID OF UNCERTAIN ETIOLOGY AND IF CLINICALLY WARRANTED A FOLLOW UP SCAN OR NUCLEAR MEDICINE STUDY COULD BE PERFORMED FOR FURTHER CHARACTERIZATION

Figure 2 One of six MARS radiology reports that was used in the study.

Radiology Report
 THYROID
 THYROID CYST
 ULTRASOUND OF THE THYROID
 Realtime ultrasonography
 solid focal nodule
 carotid artery
 Nuclear Medicine scan of the thyroid
 HEMORRHAGE INTO A CYST

Figure 3 A list of the phrases representing concepts for the report in Figure 2 that might plausibly be used to form a Medline search that is specific to this patient.

Aspiration
 Aspirations (Psychology)
 Carotid Arteries
 Diagnosis <1>
 diagnosis <2>
 etiology
 Evaluation
 Goiter, Nodular
 History <1>
 history <2>
 Measures
 Nuclear Medicine
 Radiology
 Rights
 Thyroid Gland
 Time Studies
 Ultrasonography

Figure 4 The MeSH terms returned by PostDoc for the report shown in Figure 2.

Adenoma
 Restriction Mapping
 Escherichia Coli
 Genotype
 Isomerases
 Magnetic Resonance Imaging
 Ultrasonic Diagnosis
 Brain Diseases
 Tomography, X-Ray Computed
 Receptors, Dopamine
 Thyroid Neoplasms
 Rats, Inbred Strains
 Benzazepines
 Goiter, Nodular
 Liver Neoplasms
 Molecular Sequence Data
 Sequence Homology, Nucleic Acid
 Thyroid Gland
 Dopamine
 Diagnosis, Differential
 Thyroid Diseases
 Ovary
 Carotid Artery Diseases
 Biopsy, Needle
 Goiter
 Cerebral Ischemia
 Thyroidectomy
 Hepatoma
 Cystic Fibrosis
 Microscopy, Electron
 Liver
 Equipment Design
 Nerve Degeneration
 Transducers
 Carotid Arteries
 SK&F-38393
 Head and Neck Neoplasms
 Cysts
 DNA
 Immunohistochemistry
 Radioimmunoassay
 Carotid Artery, Internal
 Corpus Striatum
 Dogs
 Blood Vessel Prosthesis
 Cerebral Angiography
 Brain
 Aphasia

Figure 5 The MeSH terms returned by Pindex for the report shown in Figure 2.

charge summaries. For each report type, six patient reports were randomly selected from MARS during the years 1992 through 1994. Thus, in total, we used 18 MARS patient reports. Information potentially identifying patients and/or physicians was systematically removed from the records before they were analyzed. Figure 2 shows one of the six radiology reports used in the study.

Before applying PostDoc and Pindex to each of the 18 reports, authors GFC and RAM performed separately the following assessment for each report:

Assessment 1. Based on the text in this report, we each circled phrases in the text that we believed represented medical concepts that could conceivably be used in doing a Medline search for this patient. In performing this task, we each tried to take the viewpoint of a Medline-experienced, but clinically novice, fourth-year medical student who is doing a Medline search after hav-

1. Adenoma
2. Aphasia
3. Aspiration
4. Aspirations (Psychology)
5. Benzazepines
6. Biopsy, Needle
7. Blood Vessel Prosthesis
8. Brain
9. Brain Diseases
10. Carotid Arteries
11. Carotid Artery Diseases
12. Carotid Artery, Internal
13. Cerebral Angiography
14. Cerebral Ischemia
15. Corpus Striatum
16. Cystic Fibrosis
17. Cysts
18. DNA
19. Diagnosis <1>
20. Diagnosis, Differential
21. Dogs
22. Dopamine
23. Equipment Design
24. Escherichia Coli
25. Evaluation
26. Genotype
27. Goiter
28. Goiter, Nodular
29. Head and Neck Neoplasms
30. Hepatoma
31. History <1>
32. Immunohistochemistry
33. Isomerases
34. Liver
35. Liver Neoplasms
36. Magnetic Resonance Imaging
37. Measures
38. Microscopy, Electron
39. Molecular Sequence Data
40. Nerve Degeneration
41. Nuclear Medicine
42. Ovary
43. Radioimmunoassay
44. Radiology
45. Rats, Inbred Strains
46. Receptors, Dopamine
47. Restriction Mapping
48. Rights
49. SK&F-38393
50. Sequence Homology, Nucleic Acid
51. Thyroid Diseases
52. Thyroid Gland
53. Thyroid Neoplasms
54. Thyroidectomy
55. Time Studies
56. Tomography, X-Ray Computed
57. Transducers
58. Ultrasonic Diagnosis
59. Ultrasonography
60. diagnosis <2>
61. etiology
62. history <2>

Figure 6 The union of the MeSH terms returned by PostDoc and Pindex for the report shown in Figure 2. This list is called the MeSH-term list. By definition this also is the output of the Union system.

ing read the given clinical report. We circled a concept when we believed that it was plausible that such a student might perform a literature search based in part on finding Medline articles that discussed the concept. The idea was to be medically inclusive of interesting concepts, but to maintain a basic sense of the type of concepts of plausible interest when searching the medical literature.

Afterwards, we compared our lists of concepts for each report. Through discussion we reached a consen-

sus that resulted in a unified list of concepts for each report. Figure 3 shows the consensus list of concepts that were developed for the radiology report shown in Figure 2. Each line contains a phrase that we believed represents a clinical concept.

Next, the Pindex and PostDoc algorithms were applied to the text of each of the 18 reports. For each report, we took the union of the MeSH terms output by Pindex and PostDoc, and we called this the MeSH-term list. The terms in the MeSH-term list for a report were numbered from 1 to n , where n is the list length. Figure 4 shows the MeSH terms output by PostDoc for the report shown in Figure 2; Figure 5 shows the output by Pindex for the same document. Figure 6 contains the MeSH-term list for the report, which was created by taking the union of the MeSH terms in Figures 4 and 5. The MeSH-term list for a given report also is by definition the output of a hybrid system that we call *Union*.

For each report, we each also performed the following assessment:

Assessment 2. For the concept phrases associated with a report, such as those phrases shown in Figure 3, authors GFC and RAM each annotated each phrase with the number of the MeSH term from the MeSH-term list that we believed adequately (and best) represented the phrase for the purposes of a Medline search. Boolean combinations of MeSH terms were used, if necessary, to represent a complex concept. In performing this task, again we each tried to take the viewpoint of a Medline-experienced fourth-year medical student who is doing a Medline search after having read the given clinical report. We used a MeSH term to represent a concept when we believed that it was plausible that such a student might use the term (alone or in combination with other terms) to represent the concept. If no such MeSH term existed in the MeSH-term list, then the phrase was not annotated.

(44 and 19)	Radiology Report
52	THYROID
(52 and 17)	THYROID CYST
(52 and 58) or (52 and 59)	ULTRASOUND OF THE THYROID
(58 or 59)	Realtime ultrasonography
28	solid focal nodule
10	carotid artery
(52 and 41)	Nuclear Medicine scan of the thyroid HEMORRHAGE INTO A CYST

Figure 7 An annotation of the concepts in Figure 3 with the MeSH terms in Figure 6.

Table 2 ■

Summary of the PostDoc (PO), Pindex (PI), and Union (UN) Results for Radiology Reports, Pathology Reports, and Discharge Summaries and for All 18 Records

	Rater 1						Rater 2						Average					
	precision			recall			precision			recall			precision			recall		
	PO	PI	UN	PO	PI	UN	PO	PI	UN	PO	PI	UN	PO	PI	UN	PO	PI	UN
Radiology reports	0.48	0.21	0.24	0.38	0.47	0.68	0.35	0.22	0.21	0.37	0.55	0.70	0.41	0.21	0.22	0.37	0.51	0.69
Pathology reports	0.52	0.16	0.22	0.49	0.43	0.68	0.34	0.15	0.17	0.45	0.45	0.66	0.43	0.15	0.19	0.47	0.44	0.67
Discharge summaries	0.48	0.15	0.19	0.36	0.38	0.62	0.46	0.16	0.19	0.33	0.41	0.63	0.47	0.16	0.19	0.34	0.39	0.62
All 18 reports	0.49	0.17	0.21	0.41	0.42	0.66	0.38	0.17	0.19	0.39	0.47	0.66	0.44	0.17	0.20	0.40	0.45	0.66

Figure 7 provides the annotation provided by one rater (RAM or GFC) for the radiology report shown in Figure 3. The numbers in the annotation denote the MeSH terms taken from the list in Figure 6. Notice that Boolean combinations of terms are used to represent some concepts. For example, the concept corresponding to the phrase *THYROID CYST* is represented by the conjunction of MeSH term 52 from the list in Figure 6 (i.e., *Thyroid Gland*) with MeSH term 17 (i.e., *Cysts*). The rater entered no annotation for the concept represented by the phrase *HEMORRHAGE INTO A CYST*. This indicates that the rater did not believe this concept was adequately represented by the list of MeSH terms in Figure 6 or any Boolean combination of these terms.

From the data provided by the annotations of each of the 18 reports by each of the two author raters, the precision and recall statistics were derived. Relative to the annotation of a given report by a particular rater, *precision* was defined as the fraction of MeSH terms output by a system for the report that were used in that annotation to represent one or more concepts. *Recall* was defined as the fraction of concepts in the annotation that were adequately represented by a MeSH term (or some Boolean combination thereof) from the MeSH terms output by a system. Based on the annotations provided by each rater, we calculated for each of PostDoc, Pindex, and Union the mean recall and precision for each of the three report types,

Table 3 ■

P Values for the Comparison of the Average Performances of the PostDoc, Pindex, and Union Systems

	Precision	Recall
Radiology reports	0.009	0.009
Pathology reports	0.003	0.019
Discharge summaries	0.003	0.009
All 18 reports	0.0001	0.0001

as well as for all 18 reports taken together. We also calculated mean recall (precision) statistics by averaging the recall (precision) statistics of the two raters.

Results

Table 2 shows the recall and precision results for the PostDoc, Pindex, and Union systems. The numerical results in the tables are rounded to two digits of accuracy. The averages shown in the tables are based on taking the average results of the two raters. The tables indicate that:

- PostDoc and Pindex each had recalls of about 40% and 50%. That is, each system outputs MeSH terms that adequately represent about 40% to 50% of the concepts in the reports.
- PostDoc had a precision of about 40% to 50%. That is, about 40% to 50% of the MeSH terms output by PostDoc were used to represent one or more concepts. Pindex had a precision of about 15% to 20%.
- When the PostDoc and Pindex outputs were taken together (to create the Union system), the recall increased to about 60% to 70%, while the precision was about 20%.

The latter result suggests that the MeSH terms output by PostDoc and Pindex were synergistic, because taken together they substantially increased recall relative to the recall of either system alone. In this experiment, MeSH terms were adequate to capture about 69% of the relevant concepts in the radiology reports, 67% of the concepts in surgical pathology reports, and 62% of the concepts in hospital discharge summaries. § While Table 2 suggests plausible similar

§ Since in this experiment the MeSH terms used to represent report concepts were limited to those terms output by the PostDoc and Pindex systems, the recall (capture) rate given here is a lower bound of the recall rate that would occur if all MeSH terms were available for coding the reports.

Table 4 ■

P Values for the Pairwise Comparisons of the Average Performances of the PostDoc, Pindex, and Union Systems

	PostDoc and Pindex		Pindex and Union		PostDoc and Union	
	Precision	Recall	Precision	Recall	Precision	Recall
Radiology reports	0.028	0.115	0.249	0.028	0.028	0.028
Pathology reports	0.028	0.917	0.028	0.028	0.028	0.043
Discharge summaries	0.027	0.345	0.028	0.028	0.028	0.028
All 18 reports	0.0002	0.267	0.001	0.0002	0.0002	0.0003

ities and differences among the three systems, the number of reports was small. Therefore, we analyzed the statistical significance of the patterns observed in that table.

Table 3 illustrates comparisons of the average recall and precision of the PostDoc, Pindex, and Union systems for each type of report. The values shown in the table were computed using the Friedman nonparametric test. The results in Table 3 indicate that it is highly likely that the average precisions of PostDoc, Pindex, and Union were not all the same. A similar conclusion holds for recall.

Table 4 gives pairwise comparisons of performances between the systems. These results were computed using the Wilcoxon signed-rank test. Table 4 indicates that the higher precision of PostDoc relative to Pindex was statistically significant, while the difference in recall between the two systems was not statistically significant. Table 4 indicates that the difference in recall between Pindex and Union was statistically significant, while the difference in precision was not significant for radiology reports. The table also indicates that the difference in recall and precision between PostDoc and Union was statistically significant. Overall, the statistical test results summarized in Tables 3 and 4 support the statistical significance of the patterns of recall and precision found in Table 2, as described previously.

Table 5 shows the average length of the MeSH-term lists output by each of the three systems, according to the type of report. The average list length output by PostDoc was consistently shorter than that of Pindex across the different report types. As expected, the list length for the Union system was longer than those of both PostDoc and Pindex.

Discussion

The results of this study suggest that the union of the outputs of PostDoc and Pindex provided significantly better coverage (recall) of clinical concepts than does

Table 5 ■

Lengths of the MeSH-term Lists Output by the Three Systems According to Report Type*

	Average length of MeSH-term list		
	PostDoc	Pindex	Union
Radiology reports	12 (5.5)	30 (14)	39 (18)
Pathology reports	15 (3.7)	37 (8.3)	48 (8.4)
Discharge summaries	30 (15)	104 (38)	125 (46)
All 18 reports	19 (12)	57 (41)	71 (48)

*In each cell, the number without parentheses represents the mean length; the number in parentheses is the associated sample standard deviation. These results are given to two digits of accuracy.

either one alone, for radiology reports, pathology reports, and discharge summaries. In general, the fraction of relevant MeSH terms generated for a report (precision) by Pindex was less than the fraction generated by PostDoc.

Both the PostDoc and the Pindex approaches could be refined to produce better performance. As noted in its description, the PostDoc algorithm employs a number of empirically derived heuristics. Many of these could be refined to improve system performance. For example, the requirement that 51% of the words in the source document must appear in the target phrase could be refined so that negations of a phrase might be recognized and excluded as possible matches (e.g., distinguishing between "Diabetes Mellitus, insulin dependent" and "Diabetes Mellitus, non-insulin dependent").

Since PostDoc was developed in 1991–92, it uses version 1.1 of the UMLS Metathesaurus. We expect that the use of the current version of the Metathesaurus, which contains more synonyms and lexical variants, would alter the performance of PostDoc. In particular, the increased coverage of the current Metathesaurus might lead to an increase in PostDoc's recall, with possibly some decrease in its precision; additional ex-

periments will be needed to know the particular effects on the performance of PostDoc.

The probability threshold at which Pindex includes terms in its output list could be increased. By altering this threshold, a user could dynamically trade off recall for precision, as needed for performing a given search.

Currently, Pindex and PostDoc output MeSH terms from among those in the entire MeSH vocabulary. A postprocessor that constrains their output to contain only *clinically* relevant MeSH terms would be likely to increase precision, while affecting recall little if any; such a restricted list of terms could be generated using semantic types in the UMLS Metathesaurus.

For the radiology and pathology reports, the length of the Union system's MeSH-term list was typically less than 50 terms. It seems plausible that a clinician user could scan such a list in short order to locate terms of interest to include in a Medline search expression. For discharge summaries, however, the Union system's MeSH-term list is about 125 terms, which may be too long to scan in an acceptable amount of time. The length of the list could be decreased by implementing the filtering methods described in the previous paragraph.

In summary, the results of this study suggest that there is a significant benefit to using a hybrid version of PostDoc and Pindex to represent in MeSH the concepts in clinical text. The coverage of clinical concepts is expected to be about 60 to 70% complete for radiology reports, pathology reports, and discharge summaries. The extent to which the MeSH vocabulary is contributing to this incomplete coverage is currently an open question. In addition, the recall and precision of PostDoc, Pindex, and Union for other types of clinical reports and uses remain to be studied.

We have reported here a formative laboratory evaluation of the abilities of three systems to identify relevant MeSH terms from clinical text. Such an in-vitro study seems appropriate at this stage, since the systems are early in their development. Fundamental issues, however, remain to be addressed by future research. We need to extend our knowledge regarding clinicians' patient-specific search needs.¹ It is important also that we understand better how to incorporate programs such as PostDoc and Pindex into comprehensive patient-specific search systems. We have much to learn from future experiments that comparatively evaluate alternative approaches to performing patient-specific searches in a clinical setting, including approaches that incorporate methods similar to those used by PostDoc and Pindex.

The authors thank Dr. John Vries for providing the Medline files that were used in training the Pindex system and for providing the MARS patient records used in this study.

References ■

- Osheroff JA, Forsythe DE, Buchanan BG, Bankowitz RA, Blumenfeld BH, Miller RA. Physicians' information needs: analysis of clinical questions posed during patient care activity. *Ann Intern Med.* 1991;14:576-81.
- Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Meth Inform Med.* 1993;32:281-91.
- Lowe HJ, Buchanan BG, Cooper GF, Vries JK. Building a medical multimedia database system to integrate clinical information: an application of high performance computing and communications technology. *Bull Med Libr Assoc.* 1995; 83:57-64.
- Miller RA, Gieszczykiewicz FM, Vries JK, Cooper GF. CHARTLINE: Providing bibliographic references relevant to patient charts using the UMLS Metathesaurus knowledge sources. *Proc Symp Comput Appl Med Care.* 1992:86-90.
- Kanter SL, Miller RA, Tan M, Schwartz J. Using PostDoc to recognize biomedical concepts in medical school curricular documents. *Bull Med Libr Assoc.* 1994;82:283-7.
- Masarie FE Jr, Miller RA. Medical subject headings and medical terminology: an analysis of terminology used in hospital charts. *Bull Med Libr Assoc.* 1987;75:89-94.
- Olson NE, Sheretz DD, Erlbaum MS. Source inversion and matching in the UMLS Metathesaurus. *Proc Symp Comput Appl Med Care.* 1990:141-5.
- Powsner SM, Miller PL. Automated online transition from the medical record to the psychiatric literature. *Meth Inform Med.* 1992;31:169-74.
- Hersh WR. Evaluation of Meta-1 for a concept-based approach to the automated indexing and retrieval of bibliographic and full-text databases. *Med Decis Making.* 1991;11 suppl:S120-S124.
- Cimino JJ, Johnson SB, Aguirre A, Roderer N, Clayton PD. The Medline button. *Proc Symp Comput Appl Med Care.* 1992:81-5.
- Vries JK, Marshalek B, D'Abarno JC, Yount RJ, Councill CD. An automated indexing system utilizing semantic net expansion. *Comput Biomed Res.* 1992;25:153-67.
- Wagner MM, Cooper GF. Evaluation of a Meta-1 automatic indexing method for medical documents. *Comput Biomed Res.* 1992;25:336-50.
- Sager N, Lyman M, Bucknall C, Nhan N, Tick LJ. Natural language processing of the representation of clinical data. *J Am Med Informat Assoc.* 1994;1:142-60.
- Friedman C, Alderson PO, Austin JHM, Cimino JJ, Johnson SB. A general natural-language text processor for clinical radiology. *J Am Med Informat Assoc.* 1994;1:161-74.
- Haug P, Koehler S, Lau LM, Wang P, Rocha R, Huff S. A natural language understanding system combining syntactic and semantic techniques. *Proc Symp Comput Appl Med Care.* 1994:247-51.
- Evans DA, Brownlow ND, Hersh WR, Campbell EM. Automating concept identification in the electronic medical record: an experiment in extracting dosage information. *Proc Fall Symposium of the American Medical Informatics Association.* 1996:388-392.
- Selden CR, Humphreys BL. Current bibliographies in medicine: Unified Medical Language Systems (UMLS). Wash-

- ington, DC: National Library of Medicine, 1997. [Ordering information: CBM number 96-8, GPO list ID: 02NLM: ZW 1 N272 no. 96-8, Superintendent of Documents, U.S. Government Printing Office, P.O. 371954, Pittsburgh, PA 15250-7954. Also available at no cost at: <http://www.nlm.nih.gov/pubs/resources.html>]
18. Evans DA. Pragmatically-structured, lexical-semantic knowledge bases for unified medical language systems. Proc Symp Comput Appl Med Care. 1988:169-73.
 19. McCray AT, Srinivasan S, Browne AC. Lexical methods for managing variation in biomedical terminologies. Proc Symp Comput Appl Med Care. 1994:235-9.
 20. Aliferis CF, Miller RA. On the heuristic nature of medical decision-support systems. Meth Inform Med. 1995;34:5-14.
 21. Schwartz JE, Miller RA, Cooper GF. Final Report on the UMLS PostDoc Project, Section of Medical Informatics, University of Pittsburgh, 1996.
 22. Yount RJ, Vries JK, Council CD. The Medical Archival System: an information retrieval system based on distributed parallel processing. Information Processing Management. 1991;27:379.

APPENDIX A

Pseudo-code Version of the PostDoc Phase 2 Algorithm

```

LOOP1: // initialize to empty
CURRENT_PHRASE = NULL STRING
CURRENT_SET = NIL
LOOP2: // get next source word (or same word if backup executed immediately before)
NEXT_WORD = GET_NEXT_WORD (SOURCE DOCUMENT phase 1 output)
IF NEXT_WORD = NIL THEN GO TO FINISH // reached end of source document
NEXT_SET = GET_INVERTED_WORD_INDEX (NEXT_WORD)
IF CURRENT_SET EQUALS NIL THEN // first time, so single word matches self
BEGIN
CURRENT_SET = NEXT_SET
CURRENT_PHRASE = CURRENT_PHRASE + NEXT_WORD
GO TO LOOP2
END
TEST_SET = SET_INTERSECTION(CURRENT_SET, NEXT_SET)
IF TEST_SET EQUALS NIL THEN // growth of phrase terminated by null intersect
BEGIN
OUTPUT (CURRENT_PHRASE, CURRENT_SET) // output matched phrase
BACKUP_ONE_WORD (SOURCE DOCUMENT) // to re-read current word next
GO TO LOOP1
END
ELSE
BEGIN
CURRENT_SET = NEXT_SET
CURRENT_PHRASE = CURRENT_PHRASE + NEXT_WORD
GO TO LOOP2
END
FINISH:
IF CURRENT_SET NOT EQUAL NIL THEN
BEGIN
OUTPUT (CURRENT_PHRASE, CURRENT_SET)
END
OUTPUT ("END OF ALGORITHM")

```

(Appendixes continue on the next page)

APPENDIX B.1

Simple Example Illustrating How Pindex is Trained

For brevity, we consider the following text as constituting the entire text of some article, although it is only a portion of an abstract that we will return to later in this appendix.

Fourteen months later mitral valve replacement (St Jude) was performed and the hemolysis ceased promptly.

After application of steps 1 through 4 in the section on the construction of associations in Pindex, we obtain the following four subtexts, where an asterisk denotes one or more stop words, which demarcate the subtexts:

* months later mitral valve replacement * st jude * performed * hemolysis ceased promptly *

Notice that parentheses are considered “stop words.”

For each of the four subtexts, we construct every phrase of word-length 1, 2, 3, and 4. Consider, for example, the subtext *months later mitral valve replacement*. From this subtext, we create the following phrases:

Phrases of length 1:

months
later
mitral
valve
replacement

Phrases of length 2:

months later
later mitral
mitral valve
valve replacement

Phrases of length 3:

months later mitral
later mitral valve
mitral valve replacement

Phrases of length 4:

months later mitral valve
later mitral valve replacement

The three-word phrase *mitral valve replacement* occurred in 88 of the articles from the six months of Medline used for constructing associations. In contrast, many of the phrases constructed (particularly nonsensical ones such as *later mitral*) occurred only rarely in the article texts.

NLM indexers assigned the following MTs to the article†:

Anemia, Hemolytic
Heart Valve Prosthesis
Mitral Valve
Mitral Valve Insufficiency
Postoperative Complications
Reoperation

A cross product is taken between the above 14 phrases and the above 6 MTs. Thus, there are a total of $14 \times 6 = 84$ phrase–MT pairs that are created for this article. For each pair, the tally for that pair is increased by one, indicating that the pair has occurred in an article. This tallying process proceeds over all articles in the six-month training set. For example, the pair (*mitral valve replacement*, Heart Valve Prosthesis) is one of the 84 phrase–MT pairs. This particular pair occurred in 60 of the training-set articles. Recall that the phrase *mitral valve replacement* occurred in 88 of the training-set articles. Thus, the frequency of Heart Valve Prosthesis appearing as an indexing term (i.e., a MT) of an article that has the phrase *mitral valve replacement* (in a title or abstract) is calculated as $60/88 = 0.68$. We use this frequency to estimate the probability that Heart Valve Prosthesis is a concept described in a free-text patient record given that the phrase *mitral valve replacement* appears in that text.

†Keep in mind that the text given above for this example is just a subset of the complete article text. Also, recall that the MTs shown do not include the check tags or similar terms.

APPENDIX B.2

Application of Pindex

In the example given we demonstrate the application of Pindex to a portion of a medical chart taken from the MARS hospital information system at the University of Pittsburgh. Some patient characteristics have been altered to maintain the anonymity of this patient.

HISTORY OF PRESENT ILLNESS

Briefly, this is a 72-year-old black female with a history of insulin-dependent diabetes mellitus, hypertension, coronary artery disease, status post myocardial infarction, Paget's disease status post left nephrectomy, and renal failure who presents with right-sided weakness.

... [text omitted here for brevity]

The patient was assessed by Physical Therapy and Occupational Therapy, a source for an embolic event was looked for. Carotid dopplers were normal and an echocardiogram showed a mild to moderately enlarged left atrium, mildly enlarged left ventricle, with abnormal left ventricular wall motion. The basal posterior wall motion was found to be akinetic or dyskinetic and the inferior wall motion hypokinetic, with overall left ventricular function moderately decreased. There was no intracardiac thrombus seen. The patient was started on aspirin.

... [text omitted here for brevity]

(Freq)	MeSH term	<--- text phrase
(1.00)	Electrocardiography	<--- echocardiogram showed
(1.00)	Postoperative Complications	<--- status post
(0.83)	Occupational Therapy	<--- occupational therapy
(0.75)	Blood Flow Velocity	<--- dopplers
(0.68)	Myocardial Infarction	<--- myocardial infarction
(0.67)	Heart Atrium	<--- enlarged left atrium
(0.67)	Echocardiography	<--- abnormal left ventricular wall
(0.67)	Myocardial Contraction	<--- overall left ventricular
(0.67)	Heart Enlargement	<--- enlarged left atrium
(0.67)	Heart Ventricle	<--- overall left ventricular
(0.64)	Coronary Disease	<--- coronary artery disease
(0.59)	Aspirin	<--- aspirin
(0.57)	Kidney Neoplasms	<--- left nephrectomy
(0.56)	Diabetes Mellitus, Insulin-Dependent	<--- insulin dependent
(0.52)	Osteitis Deformans	<--- pagets
(0.51)	Hypertension	<--- hypertension
(0.51)	Physical Therapy	<--- physical therapy
(0.50)	Insulin	<--- insulin
(0.50)	Regional Blood Flow	<--- dopplers
(0.50)	Ultrasonic Diagnosis	<--- dopplers
(0.50)	Pulsatile Flow	<--- dopplers
(0.50)	Glaucoma	<--- dopplers
(0.50)	Ophthalmic Artery	<--- dopplers
(0.45)	Nephrectomy	<--- nephrectomy
(0.43)	Review of Reported Cases	<--- left nephrectomy
(0.40)	Kidney Failure, Chronic	<--- renal failure
(0.39)	Occupational Diseases	<--- occupational
(0.36)	Blacks	<--- black female
(0.32)	Dyskinesia, Drug-Induced	<--- dyskinetic
(0.32)	Diabetes Mellitus, Non-Insulin-Dependent	<--- dependent diabetes
(0.31)	Heart	<--- left ventricular wall
(0.29)	Tuberous Sclerosis	<--- left nephrectomy
(0.29)	Carcinoma, Renal Cell	<--- left nephrectomy
(0.29)	Pyelonephritis	<--- left nephrectomy
(0.27)	Age Factors	<--- black female
(0.27)	Sex Factors	<--- black female
(0.27)	Rats	<--- basal
(0.25)	Blood Glucose	<--- insulin dependent
(0.25)	Coronary Vessels	<--- coronary artery
(0.25)	Anemia, Sickle Cell	<--- year old black
(0.25)	Adrenergic Beta Receptor Blockaders	<--- post myocardial
(0.25)	Myocardium	<--- left ventricular wall
(0.24)	Diabetes Mellitus	<--- diabetes mellitus
(0.24)	Kidney	<--- renal
(0.24)	Kidney Failure, Acute	<--- renal failure
(0.24)	Carotid Artery Diseases	<--- carotid
(0.24)	Hemodynamics	<--- ventricular function
(0.24)	Prognosis	<--- disease status
(0.24)	Antigens, Tumor-Associated, Carbohydrate	<--- disease status

We restricted Pindex in this example to generate up to 50 MTs. After 20 seconds of processing, Pindex generated the MTs in the list shown above. The text to the right of the arrow is the phrase that invoked the MT. The MT is shown just to the left of the arrow. The frequency of the MT occurrence given the phrase is shown in the column on the far left.