Automated Extraction of Medical Knowledge from Medline Citations

James J. Cimino, M.D., Laurie J. Mallon, B.A., and G. Octo Barnett, M.D.

Laboratory of Computer Science, Massachusetts General Hospital
Boston, Massachusetts 02114

The Medline database consists of over six million citations to the medical literature, indexed by the National Library of Medicine with the use of Medical Subject Headings (MeSH) and Subheadings. We propose that analysis of MeSH Headings and Subheadings in Medline citations will reveal the interrelationships among medical concepts described in the original articles. We have developed a rule-based system which postulates relationships based on the co-occurrence of MeSH Headings in Medline citations. At present, the rule base consists of 504 rules which propose 57 rélationships. When this rule base was applied to a test set of 673 citations, 93% of the proposed relationships were determined to be correct (96%, after correction of a transcription error in the rule base). We believe this approach has great potential, both for assisting acquisition of medical knowledge and for improving the quality of Medline retrievals.

Introduction

There is an explosive growth of knowledge in the medical literature. The importance of accessibility of this information has been discussed[1], but the rapid increase in medical knowledge makes retrieval an elusive, ever-receding goal.

Many schemes have been developed for accessing medical information; the most notable and extensive of these is the Medline database[2], maintained by the National Library of Medicine (NLM), which contains citations to over six million medical journal articles. Each citation includes indexing terms taken from the Medical Subject Headings (MeSH) which reflect the content of the article. For example, an article about the use of propranolol to treat patients with myocardial infarction would be indexed with the MeSH Headings PROPRANOLOL and MYOCARDIAL INFARCTION. Indexers make use of Subheadings to modify the MeSH Headings. Thus, PROPRANOLOL might be modified by THERAPEUTIC USE, while MYOCARDIAL INFARCTION would be modified by DRUG THERAPY. To some extent, the Subheading reflect the interrelationships among medical concepts discussed in the article.

Other workers have examined the use of Subheading combinations to facilitate retrieval of articles which discuss a precise relationship between specific terms[3]. We hypothesize that the Subheading combinations which appear in a Medline citation can be used to infer the concept interrelationships discussed in the cited article, providing a means for automatically obtaining medical knowledge. We have developed a rule-based system, coupled with an automated literature search-and-retrieval system, to test this hypothesis by postulating interrelationships between medical concepts. We believe that this process is capable of distilling knowledge in a way that will be useful for many medical applications. Furthermore, if the rule-base can provide insight into the implied relationships in Medline citations, there exists the potential for improving Medline retrievals.

Methods

We compared the Heading/Subheading combinations appearing in indexed citations with the relationships described in the cited articles. We postulated that co-occurrence of two Heading/Subheading combinations was sufficient to imply the relationship described in the source article. These postulated rules were tested, refined, verified and evaluated.

Definitions

We focused our interest on five classes of medical "concepts": <Anatomic Site>, <Biologic Process>, <Chemical>, <Disease>, and <Procedure>. We concerned ourselves solely with the MeSH Headings of these five classes of concepts and the relationships among them.

A *rule* is defined as an expression of some co-occurrence of MeSH Headings and Subheadings which, when satisfied by some citation, postulates a *relationship*. A *rule* is satisfied when some *target* concept (the target of a literature search) of a particular class of concepts, coupled with a specific Subheading (the *target subheading*), and an *object* concept of a particular class of concepts, coupled with a specific (usually different) Subheading (the *object subheading*) appear together in a citation. For example, if some *rule* exists with <Disease> as the *target*, "DRUG THERAPY" as the *target subheading*, <Chemical> as the *object*, "THERAPEUTIC USE" as the *object subheading*, and "<*target*> is treated by <*object*>" as the *relationship*, then the above example citation would generate the *relationship*: "MYOCARDIAL INFARCTION is treated by PROPRANOLOL". Note that some other *rule* might exist (with the *target* <Chemical> and the *object* <Disease>) such that the converse *relationship* would be postulated: "PROPRANOLOL treats MYOCARDIAL INFARCTION". **Figure 1** shows two sample *rules*, a fictitious citation, and the *relationships* postulated by the application of the *rules* to the citation.

Establishing an Initial Rule Table

We performed two types of literature retrievals for use in our initial *rule* creation: one type of search was performed using specific *target* Headings, independent of Subheading (the initial data set), while another type of search was made using specific *target subheadings*, independent of the actual *target* (the supplemental data set). The searches were performed on Medline using MicroMeSH[4] in conjunction with a search engine[5]. All searches were performed against the most recent Medline file; some searches were broadened to include back files in order to provide substantial numbers of citations.

We noted the co-occurrence of a *target* and *target subheading* with other Headings (*objects*) and Subheadings (*object subheadings*) in each citation. One of us (JJC) examined the title (and where necessary, the abstract and original article) to establish what association, if any, was present in the article between the *target*

0195-4210/88/0000/0180$01.00 © 1988 SCAMC, Inc.

and the *object*. When a *relationship* was found, the general case was postulated that any Heading in the same class as the *target* (coupled to the *target subheading*) would have the same *relationship* with any Heading in the same class as the *object* (coupled to the *object subheading*), when these two Heading/Subheading combinations appear in the same citation. This contextual examination of combinations generated a *Rule Table*, where each entry in the table consisted of a *target* class, a *target subheading*, an *object* class, an *object subheading*, and a *relationship*. A *rule* generates a single *relationship*; a *relationship* might be generated by many *rules*.

Modification of the Rules

The *Rule Table* was applied to the citations in the initial data set. The resulting *relationships* were compared to the source citations and were judged as consistent or inconsistent, based on the associations apparent in the title (and, where necessary, the abstract). If a *relationship* was judged to be inconsistent, the responsible *rule* was identified as unreliable. Several restrictions were applied to improve specificity. First, citations were only considered when the *target/target subheading* appeared as a major descriptor (Medline assigns index Headings as either major or minor). Unreliable *rules* were either modified to restrict the *object/object subheading* to the major descriptors as well, or they were discarded.

Verification of the Modified Rule Table

The modified *Rule Table* was reapplied to the initial data set, resulting in a collection of concepts and interrelationships. Each of the proposed *relationships* involving one of the four original *target* Headings was examined (as was done in the modification phase) to verify that they were consistent with the cited articles.

Evaluation of the Rule Table

To evaluate our *Rule Table* we performed an additional literature search, specifying two new MeSH Headings as *targets*, producing a test data set. The *Rule Table* was applied to the test data set, and the postulated *relationships* were examined.

Results

Establishing an Initial Rule Table

We chose four MeSH Headings (two < Diseases > and two < Procedures >) as our initial *targets*: Myocardial Infarction, Syncope, Heart Auscultation and Angiocardiography. A total of 1963 citations were retrieved and formed the initial data set.

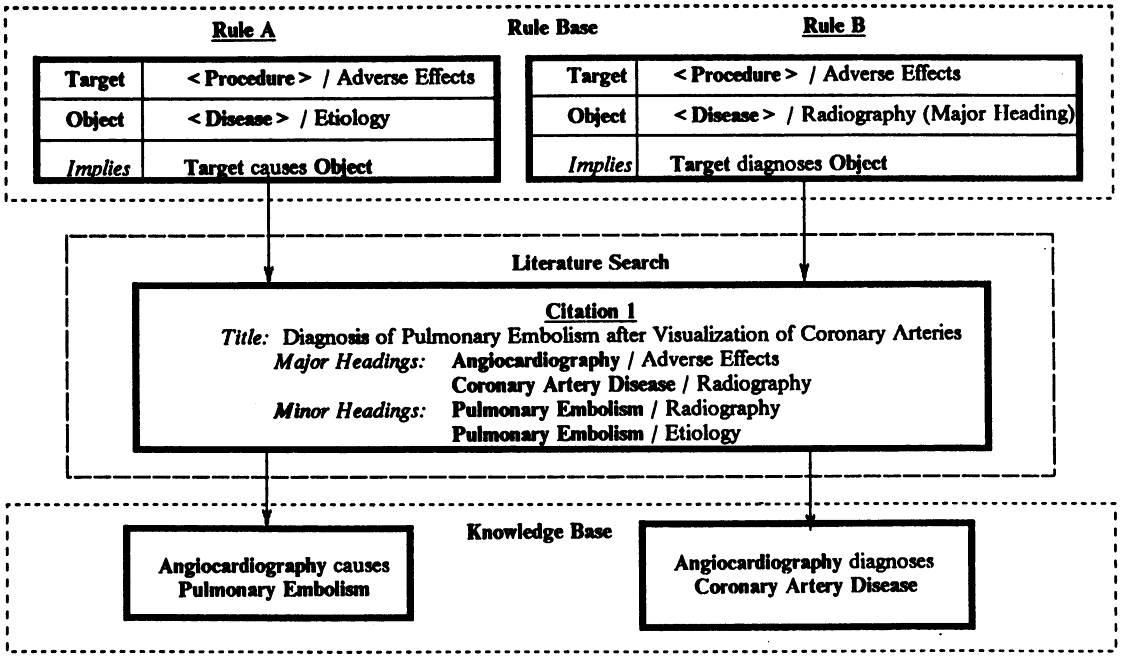
Rule A
Rule Base
Rule B
Target
<Procedure> / Adverse Effects
Object
<Disease> / Etiology
Implies
Target causes Object
Target
<Procedure> / Adverse Effects
Object
<Disease> / Radiography (Major Heading)
Implies
Target diagnoses Object
Literature Search
Citation 1
Title: Diagnosis of Pulmonary Embolism after Visualization of Coronary Arteries
Major Headings: Angiocardiography / Adverse Effects
Coronary Artery Disease / Radiography
Minor Headings: Pulmonary Embolism / Radiography
Pulmonary Embolism / Etiology
Knowledge Base
Angiocardiography causes Pulmonary Embolism
Angiocardiography diagnoses Coronary Artery Disease

Figure 1: Example of program operation

Each *rule* in the *Rule Table* is applied to every possible combination of a Major Heading (as a *target*) with every other Major or Minor Heading (as an *object*), for every citation in the data set. If the conditions of a *rule* are satisfied by the *target* and *object*, the *relationship* for that *rule* is "instantiated" with the *target* and *object* and the result is proposed.

In this example, a *Rule Table* consisting of two *rules* is applied to a fictitious citation. When "**Angiocardiography**/Adverse Effects" is the *target* and "**Pulmonary Embolism**/Etiology" is the *object*, Rule A generates the *relationship* "**Angiocardiography** causes **Pulmonary Embolism**". Similarly, when "**Angiocardiography**/Adverse Effects" is the *target* and "**Coronary Artery Disease**/Radiography" is the *object*, Rule B generates the *relationship* "**Angiocardiography** diagnoses **Coronary Artery Disease**". Note that Rule B seems counterintuitive; nevertheless, this was the relationship found to be present through examination of citations. This is largely due to the fact that Medline does not index procedures with the Subheadings such as "Diagnostic Use".

Note that the *relationship* "**Angiocardiography** diagnoses **Pulmonary Embolism**" is not proposed when "**Angiocardiography**/Adverse Effects" is the *target* and "**Coronary Artery Disease**/Radiography" is the *object*, because Rule B requires that the *object* be a major heading (see "Modifying the Rule Table" under "Methods" in the text).

Aortic Valve Stenosis

Caused by these Diseases: 7
Caused by disease of these Anatomic Sites: 10
Caused by disease affecting these Chemicals: 2
Caused by these Procedures: 3
Caused by procedure on these Anatomic Sites: 1
Caused by procedure treating these Diseases: 7
Caused by chemical affecting Anatomic Sites: 1
Affects these Anatomic Sites: 12
Affects these Chemicals: 6
Causes blood changes of these Chemicals: 6
Causes these Diseases: 13
Diagnosed by these Procedures: 7
Diagnosed by procedure with these Chemicals: 1
Treated by these Procedures: 7
Treated by these Chemicals: 2
Treated by procedure causing these Diseases: 9
Treated by procedure affecting Anatomic Sites: 8
Other Diseases treated by same procedure: 18
Other Diseases treated by same chemical: 1
Related Anatomic Sites: 1
Related Chemicals: 1
Related Diseases: 68

Echocardiography

Diagnoses these Diseases: 77
Diagnoses these Diseases as part of procedure: 3
Performed on these Anatomic Sites: 32
Prevents/controls these Diseases: 2
Treats these Diseases: 8
Causes these Diseases: 9
Chemical parts: 12
Chemical part affects these Anatomic Sites: 3
Chemical part affects these Processes: 10
Related to these Procedures: 24

Table 1: Results of Evaluation of the Test Data Set

This table shows the types of *relationships* generated by applying the modified *Rule Table* to the test data set. The number of MeSH terms under each *relationship* is given. Samples of the actual terms appear in **Table 5**.

Additional searches were performed to retrieve twenty citations for each of fourteen Subheadings associated with disease Headings (420 citations), to form the supplemental data set. Through manual examination of the indexed Headings in each of these citations we were able to detect 565 preliminary *rules*, 319 with < Disease > *targets* and 246 with < Procedure > *targets*. These *rules* were capable of proposing a total of 57 different *relationship* types.

Modification of the Rules

The initial *Rule Table* was applied to the citations in the initial data set, generating *relationships* between each of the four *target* Headings and any co-occurring Headings from any of the classes of interest. This resulted in four sets of related Headings, one for each *target*. When these sets were examined, it became clear that some of the *relationships* were spurious. In each case, the citation and *rule* responsible were identified and the conditions of the *rule* were modified such that they would no longer be satisfied by the citation. In 88 cases where the citation generated an incorrect *relationship*, the *object* appeared as a minor descriptor. In these cases, the *rule* was modified so that it would only consider citations where the *object* is a major descriptor. In 61 additional cases, the inconsistency could not be corrected by this simple restriction of the *rule* conditions, so the *rule* was deleted from the *Rule Table*. For example, the co-occurrence of < Procedure > /Adverse Effects and < Disease > /Complications (both as major descriptors) was first postulated (based on the title and abstract of a citation) to indicate that the procedure *caused* the disease. Other citations were found (with the same two Heading/Subheading combinations as major descriptors) where the actual relationship was that the disease was *treated* by the procedure, and that the procedure was causing some other disease. This discrepancy rendered the *rule* invalid.

The modified *Rule Table* consists of 504 *rules* (283 with < Disease > *targets* and 221 with < Procedure > *targets*), expanded to include the "major descriptor *objects* only" indicator.

Verification of the Modified Rule Table

When the modified *Rule Table* was applied to the citations in the initial data set, a total of 885 MeSH Headings were identified as valid *targets* or *objects* of at least one satisfied *rule*, producing a total of 6586 *relationships*. The four initial search *targets* were related to other Headings as follows: Angiocardiography, 148 Headings through 10 different *relationships*; Heart Auscultation, 63 Headings through 8 different *relationships*; Myocardial Infarction, 310 Headings through 33 different *relationships*; and Syncope, 348 Headings through 29 different *relationships*. Not surprisingly, the *relationships* postulated were all consistent with the original articles, since this same set of data was previously used to detect inconsistent *rules*.

Evaluation of the Rule Set

To test our modified *Rule Table*, we searched the current Medline file for citations to all English-language articles containing either Echocardiography (with one of the seven < Procedure > *target subheadings*) or Aortic Valve Stenosis (with one of the 19 < Disease > *target subheadings*). This yielded 673 unique citations which formed our test data set. The modified *Rule Table* was applied to the test data set and 286 concepts were identified, with 2795 *relationships* postulated among them, including 180 *relationships* (10 different types of *relationships*) for Echocardiography and 191 *relationships* (22 different types of *relationships*) for Aortic Valve Stenosis. The types of *relationships* are shown in **Table 1** and samples of the actual *relationships* are shown in **Table 2**.

Each of the *relationships* proposed for Echocardiography and Aortic Valve Stenosis was compared to the associations found in the original citations in the test data set. Of the 180 *relationships* proposed for Echocardiography, 14 (8%) were found to be erroneous, while 166 (92%) were judged to be consistent with the original citations. Of the 191 *relationships* proposed for Aortic Valve Stenosis, 13 (7%) were found to be erroneous, while 178 (93%) were judged to be consistent with the citations. Overall, the program found 27 incorrect and 344 correct *relationships* out of a total of 371, for a success rate of 93%.

The reasons for failure fell into three categories, examples of which may be found in **Table 2**. Ten cases were due to an incorrect *relationship* associated with a single *rule* in the *Rule Table* (due to a transcription error); as a result, diseases were categorized as being *caused* by a procedure, when they should have been categorized as *diagnosed* by the procedure (had this error not occurred, the program's accuracy would have risen to 96%). The remaining 17 cases were due to twelve *rules* which, although consistent in the initial data set, were found to be inconsistent when applied to the test data set. Ten of these *rules* (14 cases) can be rendered consistent by restricting their conditions such that the *object* is required to be a Major Heading (they are will remain consistent for the initial data set as well, but may generate fewer *relationships*). The last two of the ten *rules* (three cases) remain inconsistent despite the restriction of the *object* to Major Headings. Overall, 13 *rules* require changes or deletions from the 504 *rules* in the modified *Rule Table*.

Many interesting *relationships* were generated by the system which at first glance seemed erroneous. Echocardiography is generally considered a non-invasive procedure and the proposition that such a test has chemical parts (see **Table 2**) appears incorrect. However, many articles in the test data set deal with the use of various intravenous drugs as part of the technique. Similarly, while the co-occurrence of Ear, External with Aortic Valve Stenosis might seem casual, and the program's proposition (that Ear, External is an anatomic site affected by Aortic Valve Stenosis) laughable, in fact the citation dealt with ear lobe changes in patients with this disease.

Many other *relationships* were quite curious. In one case, the program proposed that Technetium is a chemical used in Heart Auscultation. This was due to the application, by an NLM indexer, of the latter Heading to an article concerning the use of a *nuclear stethoscope* (a non-auscultatory procedure). In another case, the system proposed that Tooth Extraction was diagnostic for Syncope. In fact, the responsible article discussed this very possibility, albeit tongue-in-cheek, by describing the occurrence and subsequent diagnosis of syncope in patients having their teeth extracted. Tooth Extraction was, in effect, a provocative test.

Discussion

The ability of a rule-based system, when applied to Medline citations, to detect valid interrelationships among medical concepts has been demonstrated. Discounting a single transcription error in 504 *rules*, the program was accurate 96% of the time. Nevertheless, there are many ways in which we may improve upon this performance.

The foci of our searches have been limited to cardiovascular diseases and procedures. It is likely that many useful *rules* remain undetected, since we have not examined diseases of other organ systems or specific etiologies (such as infectious diseases), nor have we looked at procedures which make use of additional techniques (such as surgical procedures). Many important Subheading-Subheading combinations (such as PATHOLOGY and CEREBROSPINAL FLUID) have not yet been examined due to the initial narrow focus of the searches.

New levels of complexity can easily be added to the present concept classes by treating concepts in different MeSH "subtrees" as members of entirely different classes. For example, some *rules* may be useful with the Infectious Diseases portion of the MeSH Disease hierarchy, but completely unreliable when applied to other classes of diseases.

Echocardiography Diagnoses these Diseases: Aneurysm, Dissecting; Aneurysm, Infected; Aortic Coarctation; Aortic Subvalvular Stenosis; Aortic Valve Insufficiency; Aortic Valve Stenosis; Aortopulmonary Septal Defect; Arteriovenous Malformations; Calcinosis; Cardiomyopathy, Congestive; Cardiomyopathy, Hypertrophic; Cerebral Embolism and Thrombosis; Cor Triatriatum; Coronary Aneurysm; Coronary Arteriosclerosis; Diabetes Mellitus, Inslin-Dependent; Ductus Arteriosus, Patent; Embolism; Embolism, Air; Endocarditis, Bacterial; Glycogenosis 2; Graft Occlusion, Vascular; Heart Rupture, Post-Infarction; Heart Septal Defects, Ventricular; Hemochromatosis; Hypertension, Pulmonary; Infant, Premature, Diseases; Lung Diseases, Obstructive; Mitral Valve Prolapse; Myxoma; Pericardial Effusion; Pericarditis, Constrictive; Pulmonary Embolism; Tetralogy of Fallot; Transposition of Great Vessels; Wolff-Parkinson-White Syndrome

Echocardiography has Chemical parts: Contrast Media; Daunorubicin; Diatrizoate; Diatrizoate Meglumine; Dipyridamole; Doxorubicin; Glucose; Halothane; Hypnotics and Sedatives; Isoflurane; Isotonic Solutions

Aortic Valve Stenosis is caused by these Diseases: Aortic Valve Insufficiency; Arteriovenous Malformations; Calcinosis; Heart Septal Defects, Ventricular; Lymphangiectasis; Transposition of Great Vessels

Aortic Valve Stenosis is caused by these Procedures: Echocardiography; Heart Catheterization; Hemodialysis

Aortic Valve Stenosis is diagnosed by these Procedures: Angioplasty, Transluminal; Dilatation; Echocardiography; Electrocardiography; Heart Catheterization; Ultrasonic Diagnosis; Vectorcardiography

Other Diseases treated by the same procedure as Aortic Valve Stenosis: Aortic Coarctation; Aortic Subvalvular Stenosis; Arterial Occlusive Diseases; Cardiomyopathy, Hypertrophic; Endocardial Fibroclastosis; Heart Defects, Congenital; Mitral Valve Stenosis; Rheumatic Heart Disease; Transposition of Great Vessels

Table 2: Sample Relationships from the Test Data Set

The above lists of MeSH terms show terms for some of the *relationship* types shown in **Table 4**. Some *relationships* are actually incorrect (that is, not truly represented in the cited articles) and are shown here in **boldface**. The *rules* which postulated the *relationships* were therefore faulty.

The *rule* which proposed that **Aortic Valve Stenosis** is caused by **Echocardiography** was actually entered incorrectly; the correct postulate would have been that **Aortic Valve Stenosis** is diagnosed by **Echocardiography**. Two different *rules* proposed that **Echocardiography** diagnoses both **Cerebral Embolism and Thrombosis** and **Diabetes Mellitus, Insulin-Dependent**. Similarly, several *rules* proposed that **Aortic Valve Stenosis** is caused by **Aortic Valve Insufficiency**, and diagnosed by both **Dilatation** and **Angioplasty, Transluminal**. These *rules* were consistent once the restriction was added that their *objects* must be major headings in the citation. The *rule* which proposed that **Daunorubicin** and **Doxorubicin** are chemical parts of **Echocardiography** was found to be generally inconsistent and was discarded.

Another potentially fruitful undertaking is the improvement of the *rules* by adding complexity to their structure. Some examples are: the *rule* might contain multiple *targets* and/or multiple *objects*, a *target* or *object* might be required to have more than one Subheading, and various exclusionary criteria could be added. The exclusionary criteria might be used to ignore "complex" citations. For example, if a citation includes many diseases with the Subheading "CHEMICALLY INDUCED" and many chemicals with the Subheading "ADVERSE EFFECTS", it might be impossible to automatically determine which chemical causes which disease.

Expanding the complexity of our concept classes and our *rule* structure should help to improve the system's specificity. No attempt has been made, however, to determine the sensitivity of our approach. Because our system relies entirely on the content of Medline, it can never serve as a sole source of medical knowledge. Despite the rapid advance of medical knowledge, much of medical knowledge remains constant and is therefore not discussed (or indexed) in the current medical literature. Our method offers no possibility of extracting knowledge about concepts which are not represented in MeSH, particularly physical findings[6] and procedures[7].

Additional caution must be exercised when interpreting the proposed relationships, as there are many opportunities for error. Like all human endeavors, the indexing of Medline citations is not an infallible process. Errors made in assigning citation headings will obviously have a direct effect on the results of our program. A more subtle problem occurs when the NLM's definition of a MeSH Heading differs from its common one (for example, the use of Heart Auscultation to describe the nuclear stethoscope).

The content of the medical literature is also a source of error. The program can not differentiate the authoritative citations from the fanciful (such as the Dental Extraction reference) or even erroneous articles. Once an article appears in the medical literature, it may be refuted, but it is not withdrawn from Medline. Our program does not attempt to determine if a relationship is been disproven by another citation. In fact, an article that states "X treats Y" and a second article that disproves "X treats Y" are likely be similarly indexed and, therefore, indistinguishable to our program.

Given the above problems of sensitivity and specificity, it is appropriate to question the utility of our system. We believe that the rule-base and *relationships* they propose will be useful in two areas.

First, the *relationships* among medical concepts can serve as a supplemental, although not exhaustive or authoritative, resource for those interested in building medical knowledge bases. Our system is particularly well-suited to the NLM-sponsored Unified Medical Language System (UMLS). One of the proposed tasks of the UMLS is to create a semantic network of selected medical terms and concepts. Since our system takes advantage of the relationships that currently exist, and seem relevant in MeSH indexing, this strategy would be of particular value in developing the UMLS since the MeSH nomenclature is expected to be one of the important components of the developing UMLS.

A second possible use for our system is to improve, in several ways, the access to the medical literature through Medline searches. The types of relationships generated from prior searches can serve as useful guides to the topics discussed in the literature (e.g., chemicals play a role in echocardiographic technique). The specific interrelations between Headings can also serve to focus a search (for example, by pointing out specific chemicals to be added to a search for Echocardiography). Finally, there is an exciting potential for improving the quality of the retrieval results by using the conditions in the *rules* as search conditions (for example, there are 37 Subheading pairs which might be used to search for articles discussing the chemical components of Echocardiography).

Conclusion

The purpose of this study was to determine the ability of a rule-based system to reconstruct, from Medline citations, the medical concept interrelationships discussed in the medical literature. Our results show that from a few hundred citations, thousands of interesting relationships can be found. There remain many facets of this approach which have yet to be explored, but the iterative process we have carried out thus far has been very encouraging.

With the current run-away growth in medical knowledge, humans are becoming more and more reliant upon computers to help in the battle to manage and access timely, up-to-date information. We believe that our approach makes a positive contribution, through the improvement of local knowledge bases, as well as through improved access to remote knowledge bases, to the current arsenal of computer tools.

Acknowledgements

This work has been supported in part through a Unified Medical Language System contract (N01-LM-8-3513) with the National Library of Medicine and through the generous contributions of the Hewlett Packard Corporation. Dr. Cimino's work has been supported by National Library of Medicine fellowship 2 T15 LM07037-4 in medical informatics.

References

[1] Covell DG, Uman GC, Manning PR. Academia and Clinic: Information Needs in Office Practice: Are They Being Met? Annals of Internal Medicine. 1895; 103(4): 596-599.

[2] National Library of Medicine, MEDLARS Management System: MEDLINE. Bethesda, Maryland.

[3] Miller PL, Barwick KW, Morrow JS, Powsner SM, Reiley CA. Semantic relationships and medical bibliographic retrieval: a preliminary assessment. Computers and Biomedical Research. 1988; 21(1): 64:77.

[4] Lowe HJ, Barnett GO. MicroMeSH: a microcomputer system for searching and exploring the National Library of Medicine's Medical Subject Headings (MeSH) Vocabulary. Proceedings of the Eleventh Annual Symposium on Computer Applications in Medical Care. IEEE Computer Society Press, 1987: 717-720.

[5] Lowe HJ, Barnett GO, Scott J, Eccles R, Foster E, Piggins J. Remote Access MicroMeSH: A microcomputer system for searching the Medline database. Proceedings of the Twelfth Annual Symposium on Computer Applications in Medical Care. IEEE Computer Society Press, 1987: *in press*.

[6] Cimino JJ. Examination and comparison of physical examination terms and structures in existing thesauri. Report to the National Library of Medicine for the Unified Medical Language System; Bethesda, Maryland; March 9, 1987.

[7] Cimino JJ. Comparison of Cardiovascular Procedures in MeSH, CPT and ICD9. Report to the National Library of Medicine for the Unified Medical Language System; Bethesda, Maryland; September 16, 1987.