The Medline Button

James J. Cimino, M.D.; Stephen B. Johnson, Ph.D.; Anthony Aguirre, M.L.S.;
Nancy Roderer, M.L.S.; and Paul D. Clayton, Ph.D.

Center for Medical Informatics
Columbia University College of Physicians and Surgeons
New York, New York 10032

We have developed a computerized method for performing bibliographic searches directly from patient data involving five steps: 1) identifying specific patient data which raises a question in the mind of the user, 2) selection (from a list of generic questions) of a small number of questions which fit the selected patient data, 3) automated translation of the patient data into appropriate terms used for bibliographic indexing, 4) conversion of the question selected by the user into a search strategy, and 5) transfer of the search strategy to a search engine for a bibliographic database. We have modified the Columbia-Presbyterian Clinical Information System to experiment with this method. The first implementation converts patient diagnoses and procedures coded in ICD9-CM into Medical Subject Headings (MeSH) and searches Medline using BRS/Onsite. Challenges include development of a useful set of generic questions and translation from ICD9-CM to MeSH using the Unified Medical Language System (UMLS ᵗᵐ).

INTRODUCTION

A goal of the Integrated Academic Information Management System (IAIMS) at the Columbia-Presbyterian Medical Center (CPMC) is to bring together clinical data, medical knowledge and expert systems to assist in patient care through "one-stop information shopping"[1]. Another goal is to support clinical decisions through integration of diverse applications to help users solve complex diagnostic and therapeutic problems[2]. Consider the following, recently-published scenario[3].

A physician is using a clinical results-display application to review patient data. During review, a question arises for which the physician requires additional information. A second application in the system is known to be a good resource for such information, so the physician invokes it, poses an appropriate query, obtains the required information and, with the question now answered, returns to the results-display application.

In this scenario, applications are brought together through a common access point (the physician's terminal or workstation), but it is the user who must perform the real integration. For the sequence to occur, the user must select an information application, must know how to evoke and use the application (including the proper format and vocabulary needed to formulate an appropriate query) and must compose and enter the information request.

We postulate that this process can be assisted through an informed approach to application integration. First, we believe that the use of clinical applications provides a setting in which information needs arise. Second, in a given clinical application setting (e.g., a particular point in a results-review application) we believe (based on our collective clinical, educational and library experience) that some finite number of general (or "generic") questions can express a large proportion of information needs which might arise. Third, the information present in the clinical application can play a role in retrieving information to answer the question, if it is translated into a form recognized by the target information source. Fourth, given a question, an information source and properly translated patient data, an appropriate query strategy can be constructed. Fifth, it should be possible to evoke the target application automatically and transfer the query, with minimal effort required of the user. The validity of these points (particularly the second and fourth) are presently unknown and require testing.

We have developed an application which uses information displayed in a clinical information system to generate a list of specific questions. The application translates a user-selected question into a query, passes it to a computer-based information source, and displays results to the user. Because the application accesses the National Library of Medicine's Medline database, and because it allows selecting choices rather than typing queries, the application is dubbed "The Medline Button".

THE SETTING

An overall description of the CPMC Clinical Information System appears elsewhere[1]. We have chosen the Admission Profile application as the first clinical application setting in which to explore our approach. This application displays a list of encounters (hospital admissions and emergency room visits) for a given patient. When one of these encounters is selected, a list is displayed showing conditions diagnosed and diagnostic/therapeutic procedures carried out during the encounter, all of which are coded in the *International Classification of Diseases, Ninth Edition, Clinical Modifications* (ICD9-CM)[4]. If an information need arises at this point, we postulate that one or more of these diagnoses and/or procedures is involved; we have modified the CPMC Admission Profile screen to allow the user to indicate items of interest. Figure 1 shows a sample screen from the modified application.

```
Name: JOHNSON, JOAN                    MRN: 0010104
               Admission Record Detail
-----------------------------------------------------
Admission Date: 11/23/90   Discharge Date:  09/19/91
Doctor: ROSEN, MORTIMER G       Discharge Summary:  N
Select Terms You Are Interested In:

         Diseases:
         658.11   PREM RUPT MEMBRAN-DELIV
    X    414.10   ANEURYSM OF HEART WALL
         V27.0    DELIVER-SINGLE LIVEBORN
         645.01   PROLONGED PREG DEL W W/O ANTEP

         Procedures:
         73.6     EPISIOTOMY
    X    99.63    CLOSED CHEST CARDIAC MASSAGE
         75.34    FETAL MONITORING NOS

-----------------------------------------------------
 Help=F1     Admission List=F3   Search MEDLINE=Enter
 Scroll Down=F8     CIS Main Menu=F12    Signoff=F11
```

Figure 1: Typical display of admission record in Admission Profile application. User has selected a disease (Aneurysm of Heart Wall) and a procedure (Closed Chest Cardiac Massage).

PROVIDING GENERIC QUESTIONS

Certainly, the kinds of questions a user might have about diseases and procedures are innumerable. For example, since there are 12,521 diseases in ICD9-CM, the user could have at least 12,521 questions about the cause of these diseases. However, if all 12,521 questions can be reduced to the question "What causes X?", where X is a disease of interest, we can reduce these questions to a single "generic" question. If the user selects "Aneurysm of Heart Wall", we can display "What causes Aneurysm of Heart Wall?"

We are analyzing user questions (posed to reference librarians), to determine a useful set of generic questions which express a majority of user questions. Analysis is being done manually by expert librarians and through natural language processing techniques. Analysis of 40 complex user questions has elicited 68 simple user questions which typically involve one or two medical terms and a relation. The relation may be between two terms (such as "Does X cause Y?") or between a single term and an unnamed second term (such as "What causes X?"). When the types of terms and relations of the 68 questions were analyzed, we found that 37 generic questions (expressed as "type - relation - type", with the second type optional) could account for all of the user questions. For each generic question, a search strategy was generated. The purpose of this paper is to describe the Medline Button, our prototype system in which generic questions are integrated into a clinical application; further description of the generic questions themselves is beyond the scope of this paper.

Generic questions relevant to diseases and procedures (18 in all) were incorporated, with their search strategies, into the CPMC Medical Entities Dictionary (MED)[5] and Metadatabase[6] (the central repository of information about the what, where, and how of storing coded patient data in the CPMC clinical database). The Medline Button selects from the Metadatabase those questions which fit user-selected clinical information. For example, if the user chooses a disease and a procedure in the Admission Profile, questions which can accommodate both are shown (Figure 2).

```
Name: JOHNSON, JOAN                    MRN: 0010104
                 Medline Questions
-----------------------------------------------------
Select the question you are interested in:      4

    1    Does Heart Massage prevent Heart Aneurysm?

    2    Does Heart Massage diagnose Heart Aneurysm?

    3    Does Heart Massage treat Heart Aneurysm?

    4    Does Heart Massage cause Heart Aneurysm?

-----------------------------------------------------
 Help=F1                      Search MEDLINE=Enter
 Scroll Down=F8    CIS Main Menu=F12   Signoff=F11
```

Figure 2: List of potentially relevant generic questions, modified to reflect the ICD9 terms selected by the user. Note that terms have been translated to MeSH equivalents. In this example, the user has selected the fourth question.

TRANSLATING PATIENT INFORMATION

The CPMC IAIMS includes BRS/Onsite with a database of over one million Medline bibliographic citations. The Medline Button invokes BRS to provide information relevant to selected questions. ICD9-CM terms could be used to search Medline in a free-text manner, but their form is usually not ideal for this purpose (see, for example, the terms listed in Figure 1). A more appropriate method of searching Medline involves the identification of Medical Subject Heading (MeSH) terms which are relevant to the selected ICD9-CM terms, and then retrieving Medline citations indexed by those MeSH terms.

The translation from ICD9-CM to MeSH is not a trivial task. Previous work has shown that the National Library of Medicine's Unified Medical Language System (UMLS^{tm})[7] can help identify associated MeSH terms for a large proportion of ICD9-CM terms[8]. However, when these associations were examined in the context of use for bibliographic searching, it was found that, in most cases, they were not sufficiently precise for translation. We therefore carried out an extensive effort to provide improved translations for ICD9-CM to MeSH, based on the UMLS. 705 terms were found to be synonymous with a single MeSH term (614 of these relationships were found in the 1991 version of the UMLS Metathesaurus), an additional 288 were found to match closely to a single MeSH term, and 308 more were found to be best translated to two or more MeSH terms, for a total of 1301 terms translated. Of those terms we did not translate, many have more general entries in either ICD9-CM or in the MED for which some translation does exist. If no MeSH translation exists for a more specific term, the translation of the more general term is used, although it is often less than ideal. When these cases are taken into account, 8873 of the total 18,631 ICD9-CM terms (48%) can now be translated to MeSH.

ICD9-CM terms are a major component of the MED[9]. The MeSH translations have been added to the MED and the Metadatabase, making conversion from an ICD9-CM term (or code) to the appropriate MeSH term(s) a table-driven operation. When generic questions are to be displayed by the Medline Button, the ICD9-CM terms are translated to MeSH and inserted into the text of the questions, which are then displayed to the user (see Figure 2). The user may then select one of the questions as relevant to his or her original information need.

CONSTRUCTING A SEARCH STRATEGY

The question selected by the user is associated with a generic search strategy. Like the question, it is filled in with the MeSH term(s) obtained by translating the selected ICD9-CM term(s). For example, the generic question "What causes X?" can be coupled to a generic search strategy such as "X with ET", where "X" is the place holder for a disease term, "ET" is the Subheading for "etiology" in the Medical Subject Headings (MeSH)[10], and "with" is a BRS search operator which specifies that the MeSH term and Subheading are to appear in the same sentence in a retrieved citation.

AUTOMATIC EVOCATION OF SOURCE

Access to IAIMS applications is accomplished using a personal computer with scripting facilities. The scripting facility is used to assist in user log-ons and application switching, including access to multiple simultaneous applications. We have modified the script to recognize when a user has selected a question in the Medline Button. It then takes control of the user's Admission Profile session, sending the application a message to display the search strategy which corresponds to the selected question. The script then logs on to Medline, transfers the search strategy and displays the results (see Figure 3). At this point, control of the Medline session is returned to the user, who can display citations, perform additional searches, or review the search strategy. When the user is finished using Medline, the script returns the user to the original Admission Profile session.

```
SEARCH   1           MEDLINE        SCREEN     1 OF 1
         1
AU  Chatson-G.  Gallagher-R.  Quahliero-D.
TI  Ventricular pseudoaneurysm associated with
    cardiopulmonary resuscitation 6 weeks after
    mitral valve replacement.
SO  Ann-Thorac-Surg.  1989 Nov.  48(5).  P 719-20.
AB  Trauma to the heart and mediastinum is associated
    with external cardiac massage. A patient had
    undergone a redo mitral valve replacement and
    experienced an uneventful postoperative course.
    During a visit to her physician 6 weeks after
    operation, she experienced ventricular
    fibrillation that required external cardiac
    massage and subsequent defibrillation. Post
    resuscitation evaluation revealed a posterior·
    pseudoaneurysm of the ventricle. This was
    repaired via a transthoracic approach with the
    use of profound hypothermia.
MJ  HEART-ANEURYSM: etiology (et). HEART-VALVE-
    PROSTHESIS. RESUSCITATION: adverse-effects (ae).
MN  HEART-ARREST: etiology (et), therapy (th).
    HEART-MASSAGE: adverse-effects (ae).  HUMAN.
```

Figure 3: The result of the search corresponding to the question selected in Figure 2.

SYSTEM STATUS

At this writing, the Medline Button is operational in a test environment. Before the system can be placed in the CPMC "production" environment, the translation of ICD9-CM must be extended to cover the entire terminology. Further work is underway as part of a UMLS development contract to determine the kinds of information needs that are likely to arise while using the Admission Profile application and to provide complete coverage of ICD9-CM-to-MeSH translation.

DISCUSSION

The Medline Button is one of our first attempts at CPMC to integrate IAIMS applications in a clinically relevant manner. There are many other clinical applications (e.g., laboratory results review) to which this approach can be extended. Additional implementations will require analysis of the information needs that can be anticipated by users of those applications and the development of a translation method from terms in those applications into terms which can be used for information retrieval. The retrieval source options should be extended to include on-line information resources in addition to Medline. The choice of information source can be made by simply assigning a source for each question in advance, but more sophisticated approaches may be useful. For example, as the UMLS Information Sources Map develops, we may find ways to use this resource to help automate information source selection.

Many researchers in medical informatics share the desire to integrate programs in ways similar to the Medline Button. A number of prototype systems have been developed and described. Powsner and Miller developed the PsychTopix system which identifies key words in psychiatric reports and matches them with predetermined topics[11]. Their search strategies are not specific to the terms used in the patient record, but rather to more general, predefined topics. Cimino and Barnett developed the Interactive Query Workstation (IQW) which scans patient records for controlled terms which can then be recognized as relevant to one or more predefined queries. The queries can then be used to construct search strategies which are applicable to the original patient record terms and are formatted for a number of different query-specific information sources[12].

These systems have much in common with the Medline Button. A distinction of our system, however, is that generic questions are being derived from an analysis of actual user questions, rather than questions that are perceived as relevant by the system designers. Others have performed analyses of physician information needs as expressed in the in-patient[13,14] and the out-patient[15,16] settings. These studies have shown that most questions can be accounted for by a small number of general categories or types. Using different sources (library reference questions) and different analytic techniques (based on library science and natural language processing), we are defining information needs that can be dealt with in a systematic fashion.

The Medline Button is fully operational: it generates questions from selected patient information, it translates the questions into Medline search strategies, and it executes the searches automatically. However, it is too early to tell if the Medline Button approach to meeting clinical information needs is effective. The system suggests questions and provides answers, but we do not yet know if they are the right questions and the appropriate answers. We are continuing to build our set of generic questions and their corresponding search strategies. Public availability of the system awaits completion of current UMLS efforts to fully integrate ICD9 with MeSH in the Metathesaurus as well as development of reliable translation algorithms. When this work has advanced sufficiently, the Medline Button will be ready to serve as a laboratory in which to test our hypotheses of clinician information needs. The lessons learned at CPMC may have broad applicability: virtually every hospital in the United States has computer applications containing ICD9-CM data and almost every hospital library conducts Medline searches.

In traditional retrieval systems, the user has the freedom to formulate questions with great precision, but must then work with the program to express the question as a search statement. The user has the right question; asking it the wrong way will produce unsatisfactory results. In systems such as the Medline Button, IQW and Psychtopix, the user works with the system to choose a question. With predetermined questions, retrieval strategies that are more sophisticated than those the user might otherwise devise can be developed in advance. The key will be to ask the right questions.

CONCLUSION

The Medline button offers a method to reduce a complex sequence of information management steps to as few as three or four button pushes. A user looking at a computer screen full of clinical data can indicate, in a general way, those data which have triggered an information need. The system interprets this as a specific request and, with the user's approval, attempts to satisfy that request. The results may satisfy some of the people some of the time. Further work is needed to learn more about how to anticipate information needs, select appropriate sources and translate information from one system for use in another. The Medline Button is a platform for such work.

ACKNOWLEDGEMENTS

References

1. Hendrickson G, Anderson RK, Clayton PD, Cimino JJ, Hripcsak G, Johnson SB, McCormack M, Sengupta S, Shea S, Sideli R, Roderer N. The Integrated Academic Information Management System At Columbia Presbyterian Medical Center. *MD Computing*, 1992; 9(1):35-42.

2. Cimino JJ, Sengupta S: IAIMS and UMLS at the Columbia-Presbyterian Medical Center. *Medical Decision Making*, 1991;11(sup.):S89-93.

3. Guyatt GH: Evidence-based medicine. *ACP Journal Club, Annals of Internal Medicine*, 1991;114(suppl.2):A-16.

4. United States National Center for Health Statistics. *International Classification of Diseases, Ninth Revision, with Clinical Modifications*, Washington, DC; 1980.

5. Cimino JJ, Hripcsak G, Johnson SB, Friedman C, Fink DJ, Clayton PD: UMLS as knowledge base - a rule-based expert system approach to controlled medical vocabulary management. In Miller RA, ed.: *Proceedings of the Fourteenth Annual SCAMC*; IEEE Computer Society Press, New York; November, 1990:175-180.

6. Johnson SB, Cimino JJ, Friedman C, Hripcsak G, Clayton PD: Using metadata to integrate medical knowledge in a clinical information system. In Miller RA, ed.: *Proceedings of the Fourteenth Annual SCAMC*; IEEE Computer Society Press, New York; November, 1990:340-344.

7. Humphreys BL: *UMLS Knowledge Sources - Second Experimental Edition Documentation*, Bethesda, Maryland: National Library of Medicine, August, 1991.

8. Cimino JJ, Sideli RV: Using the UMLS to bring the library to the bedside. *Medical Decision Making*, 1991; 11(sup.):S116-S120.

9. Cimino JJ, Barrows RC, Allen B: Adapting ICD9-CM for clinical decision support. In: Musen MA, Cox J (eds.): *Proceedings of the AMIA Spring Congress*, May 7-9, 1992, Portland, Oregon; 34.

10. National Library of Medicine. *Medical Subject Headings*, Bethesda, MD, 1990 (NTIS NLM-MED-90-01).

11. Powsner SM, Miller PL: Linking bibliographic retrieval to clinical reports: PsychTopix. In Kingsland LC (ed.): *Proceedings of the Thirteenth Annual SCAMC*, IEEE Computer Society Press, New York, 1989:431-435.

12. Cimino C, Barnett GO: Standardizing access to computer-based medical resources. In Miller RA (ed): *Proceedings of the Fourteenth Annual SCAMC*, IEEE Computer Society Press, New York, 1990:33-37.

13. Osheroff JA, Forsythe DE, Buchanan BG, Bankowitz RA, Blumenfeld BH, Miller RA: Physicians' information needs: analysis of questions posed during clinical teaching. *Annals Internal Medicine*, 1991; 114(7):576-581.

14. Forsythe DE, Buchanan BG, Osheroff JA, Miller RA: Expanding the concept of medical information: An observational study of physicians' information needs. *Computers and Biomedical Research*, 1992; 25(2):181-200.

15. Cimino C, Barnett GO: Analysis of physician questions in an ambulatory care setting. In Clayton PD, ed.: *Proceedings of the Fifteenth Annual SCAMC*; McGraw-Hill, New York, 1991:995-999.

16. Covell DG, Uman GC, Manning PR: Information needs in the office practice: Are they being met? *Annals of Internal Medicine*. 1985; 103(4):596-599.