SmartQuery: Context-Sensitive Links to Medical Knowledge Sources from the Electronic Patient Record

Susan L. Price, MD, MS, William R. Hersh, MD, Daniel D. Olson, Peter J. Embi, MD
Division of Medical Informatics & Outcomes Research
Oregon Health & Science University, Portland, OR

ABSTRACT

Physicians have many unmet information needs that arise in the course of patient care. Many clinical questions could potentially be answered by streamlined access to medical literature, textbooks, and clinical guidelines in the context of the electronic medical record. We designed and implemented SmartQuery, a prototype application to provide context-sensitive links from an electronic patient record to relevant medical knowledge sources, then performed a preliminary user evaluation. Our results suggest that such an application may be clinically useful, and provide some insight into problems and priorities for future development.

INTRODUCTION

Many clinical questions arise in the course of patient care. Although frequency estimates vary, one reasonable estimate, for outpatient primary care, is two questions are identified for every three patient encounters.[1, 2] Of these questions, only about 30% are immediately answered.[1-3] Electronic information sources, particularly bibliographic access to journal literature, are infrequently used.[4, 5] Yet a significant proportion of these questions can be answered using electronic knowledge sources.[6, 7]

Although physicians now have access to a wide variety of electronic resources, if resources cannot be searched quickly and successfully, with little specialized training, the potential benefits may go unrealized. Possible interventions to promote successful searching include: making resources immediately available when the question arises, using patient data to anticipate the information need, aiding in query formulation and submission, and providing a single point of access to multiple resources. Linking the electronic medical record (EMR) directly to electronic resources can aid in achieving these goals.[8]

Project vision and objectives

The overall goal of SmartQuery is to provide context-sensitive links from the EMR to relevant medical knowledge sources. More specifically, we want to help a clinician find answers to clinical questions that arise while reviewing a patient's electronic data without having to leave the EMR and open a new application. Furthermore, we want to facilitate searching by automatically aiding query generation and query submission. Since patient problems are often complex and may span multiple interacting problems, or involve multiple types of information such as diagnoses and lab test results, we want the system to aid in forming complex queries, arising from information presented in various parts of the EMR. We want to provide access to different types of information sources, such as electronic textbooks, research articles, and clinical guidelines for any given query, yet limit information overload. Finally, we want the user to be able to read selected items in detail.

SYSTEM DESCRIPTION

Oregon Health & Science University (OHSU) uses the Lifetime Clinical Record (LCR), a mainframe-based reporting system from Shared Medical Systems (SMS, now Siemens), for its inpatient clinical care. SMS recently developed NetAccess, a Web-based front end to the LCR. SmartQuery, an application that provides context-sensitive links from specific patient data viewable in NetAccess to relevant medical knowledge resources, was developed as a joint development project with SMS and OHSU.

Extracting patient-specific data

Data specific to a single patient are the input for a collection of Perl CGI scripts that form the nucleus of SmartQuery. As soon as a patient's EMR is accessed, a base set of MeSH terms is collected from that patient's diagnosis list. Since the OHSU system does not contain an electronic problem list, we use ICD-9 codes as a proxy for a real patient problem list. The ICD-9 codes are translated to corresponding MeSH terms by SAPHIRE,[9] an application that can translate among the various vocabularies that are represented in the UMLS. For each ICD-9 code, SAPHIRE attempts to return corresponding MeSH terms, although in some cases none are found.

Interaction with the user

Additional terms can be added by the user while viewing the EMR. SmartQuery adds a collection of

buttons and checkboxes to the patient data displays in NetAccess. Checkboxes appear next to each lab test identifier, and above displays of dictated reports. To use SmartQuery, the user checks the boxes next to items relevant to his or her question, then clicks an *Add* button, which causes a MeSH term corresponding to the data underlying the displayed information to be added to the list of MeSH terms created from the ICD-9 codes. When the *Go* button is clicked, the MeSH terms collected from the patient's ICD-9 code list and from the user's entries via the *Add* buttons are presented as a list of checkboxes. The user may also enter his or her own terms via a textbox. At the top of this frame are a series of checkboxes for the various information sources available. To send a query, the user need only check the terms of interest and the information sources he or she wishes to query. (Figure 1)

Because some resources, especially MEDLINE and Harrison's Online, can return an overwhelming number of search results, only the top five results from each resource are displayed. The results are displayed with hyperlinks so that the user can view each result with a single mouse click. (Figure 2)

Processing of dictated (text) reports

Because a list of ICD-9 codes may not reflect problems new to the current admission, the most recent History & Physical (H&P) dictation is also used as a source of MeSH terms. We took advantage of the fact that H&P's are dictated in a fairly predictable format and only analyzed selected sections related to chief complaint, assessment, and plan since those sections are most likely to contain information about significant diagnostic and therapeutic issues. The text in each section was first pre-processed to remove section headings, expand or remove many common abbreviations, and eliminate some stray punctuation characters, ordinal numbers, and extra spaces. Each noun phrase in the resulting text was sent to SAPHIRE to get a list of associated MeSH terms. The resulting list was further filtered by removing terms from a list of commonly yielded but non-useful terms (such as "Normal"), and terms belonging to certain semantic types (such as "Biomedical Occupation or Discipline").

Query Processing

SmartQuery currently accesses five representative resources: MEDLINE, Best Evidence, the online version of Harrison's Principles of Internal Medicine textbook,[10] the National Guidelines Clearinghouse,[11] and CliniWeb,[12] a collection of human-reviewed and MeSH-indexed Web pages. Queries appropriate to each of the selected resources are formulated from the terms selected by the user. Each resource requires a different query formulation process. For example, MEDLINE and Best Evidence, are queried using MeSH terms and Boolean operators. Harrison's Online has its own full text search function and is not indexed with MeSH. Nevertheless, we used MeSH terms to query Harrison's Online except for some terms derived from lab reports.

Processing terms from lab reports

Translating the EMR data from lab reports to query terms posed two challenges. First, an abnormal value contains much more information about the patient than merely the fact that a lab test was done. Second, lab results are reported using names that do not readily form sensible queries, such as "Meas ICA, Wh B" for a calcium test. We manually developed a table for common lab tests that maps lab test identifiers to useful query terms, including terms to be used when a flag indicating abnormally high or low values is set in the LCR. For example, if the checkbox next to a patient's "Meas ICA, Wh B" results is selected, SmartQuery detects whether the most recent result was normal, high, or low, then

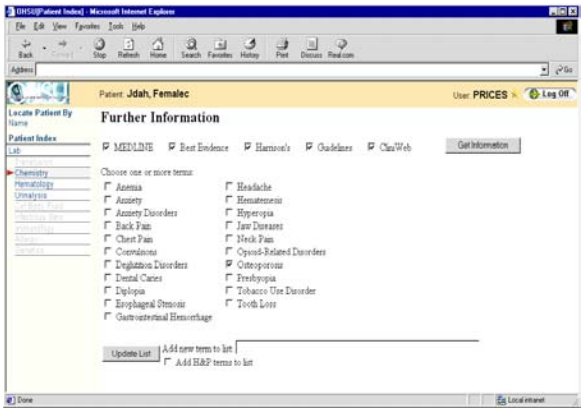

Figure 1: Query terms and resource options displayed by SmartQuery

Figure 2: Results of a search on osteoporosis

offers the term "Calcium," "Hypercalcemia," or "Hypocalcemia."

PRELIMINARY EVALUATION

Methods

We undertook a formative evaluation to assess overall usefulness and usefulness of specific features, and to guide future development of the system. We used a convenience sample of 11 OHSU residents, 7 from internal medicine and 4 from family practice.

After a brief demonstration of how to use SmartQuery, the participants performed three tasks that simulate ways SmartQuery might be used in clinical practice. Each task was embedded in a clinical scenario. Briefly, the three tasks consisted of (1) formulating an approach to caring for a complex patient (2) finding an answer to a specific question related to a patient's treatment (3) preparing for a morning report discussion about a question related to interactions between two specific patient problems. Fifteen minutes was allotted for each task, during which residents briefly recorded the answers they found and the resources they used. Upon task completion, subjects were given a survey to elicit their opinions on the usability and usefulness of SmartQuery. The scenarios and questionnaire were iteratively developed using four physicians who were also medical informatics graduate students as pilot subjects.

Questions assessed the subjects' overall impression of SmartQuery, the usefulness of specific features, and how the subjects thought they would use SmartQuery *or a similar, improved version*. Subjects were asked to evaluate the concept, not just its preliminary implementation. Questions about how it would be used solicited as many answers as applied, then asked the respondent to choose the single option that best described how it would be most valuable or used most often. Additional questions solicited open-ended feedback about SmartQuery.

Questions	Number of responses				
	SA	A	N	D	SD
Overall I liked using SQ	1	9	1	0	0
Easy to learn to use	3	7	1	0	0
Easier to find information I wanted than with the methods I currently use	0	3	4	4	0

Table 1: Responses to general questions about the usability of SmartQuery (SQ). SA=strongly agree, A=agree, N=neutral, D=disagree, SD=strongly disagree

Questions	Number of Responses		
	VU	SU	NU
How useful is the SQ feature that allows you to query multiple resources from a single interface?	7	3	1
How useful is the SQ feature that automatically generates query words?	5	5	1

Table 2: Users' responses to questions about the usefulness of specific features. VU=very useful, SU=somewhat useful, NU=not very useful

Results

As shown in Tables 1 and 2, the residents generally liked SmartQuery, learned to use it easily, and were moderately positive about it being easier to find information using SmartQuery than using customary methods. Being able to query multiple resources from a single interface was deemed more useful than the automatic generation of query words applicable to a particular patient, although nearly all thought each feature was at least somewhat useful. Tables 3 and 4 display the responses to questions about specific features and uses of SmartQuery. The first column shows the number of positive responses when residents were allowed to select all appropriate options. The second column shows the number of residents selecting a particular response as the **most** appropriate option.

Questions	Options	
	All	Best
Features of SQ that would be valuable		
Direct access from EMR to guidelines	11	2
Direct access from EMR to electronic textbooks	10	6
Direct access from EMR to automatic MEDLINE queries	10	0
Automatic generation of queries	6	0
Other	5	3
General uses of SQ (one resident chose two best options)		
Be certain all appropriate treatment strategies considered	11	0
View current guidelines; affirm that care is appropriate	10	3.5
Look up specific questions about specific patients	10	4
Generate more complete differential diagnosis	9	2.5
Access most recent literature on subject	7	0
Do research for a talk, presentation, or paper	5	0
Learn more; not related to a particular patient	3	1
Patient-specific uses of SmartQuery		
Help make treatment decisions	10	7
Help learn more about patient's disease (s)	10	3
Help make diagnostic decisions	10	0
Help learn about diagnostic and treatment strategies for a patient	8	1
Help prepare a talk, presentation or paper	7	0

Table 3: Number of positive responses to questions about features and uses of SmartQuery when asked to select all appropriate options and to select the single most appropriate option.

Questions	Options	
	All	Best
Stages in patient care		
Upon admission, after seeing the patient	10	8
When unexpected events occur	9	1
When unexpected diagnostic results occur	9	1
Upon admission, before seeing the patient	8	1
Preparing for morning report	8	0
Preparing for daily work rounds	3	0
Sections of the EMR to initiate a query		
Lab	9	2
Most recent History & Physical	8	6
Microbiology	8	0
Other dictated report	6	1
Patient problem list (ICD-9 codes)	6	1
Pathology	3	1
Other results	2	0
X-ray	1	0
Other (Patient medication list)	1	0

Table 4: Number of positive responses to questions about how and when SmartQuery might be used. Respondents asked to select all appropriate options and to select the single most appropriate option.

Eight of 11 residents predicted they would use SmartQuery at least once a day. Most indicated that we were only partially successful with generating all the necessary query words and a majority indicated that at least some of the suggested query words were inappropriate. The responses to open ended questions emphasized the residents' desire for fast, efficient access to resources and included suggestions for additional resources they would like added.

DISCUSSION

Residents who evaluated the prototype were enthusiastic about the concept, although they found limitations in its current embodiment. The success of such a system will depend upon whether it can save time for clinicians. Providing a single portal from which to submit a query to multiple resources is a significant step in improving efficiency. However, if the desired information is not retrieved, then a speedy retrieval is meaningless. The quality of the retrieval set is critical and depends upon three factors:

1. The resources being searched must contain the desired information. The residents in our study felt guidelines, textbooks, and journal articles were valuable. In addition they suggested we add drug information resources and a source of concise evidence-based summaries.
2. Patient-specific information must be converted into a list of query terms that is both comprehensive and appropriate.
3. Once terms are selected, they must be incorporated into queries, customized for each interface, that will retrieve desired information. When excess documents are returned, appropriate subsets must be selected for display.

We used a combination of ICD-9 codes, limited natural language processing of dictated text, and custom-created tables to extract information and map it to MeSH terms. Our term lists were not completely satisfactory for several reasons. Diseases and concepts represented in the EMR may have been missed due to inadequate representation of patient problems by ICD-9 codes,[13] or because complete matching between ICD-9 codes and MeSH terms is problematic.[14] Additionally, failure to match free text to MeSH terms may have been caused by failures in preprocessing, inadequacies in the SAPHIRE algorithm, or deficiencies in the completeness and granularity of MeSH itself. Clinical information that was not in the EMR, whether it was a new problem, a newly considered diagnosis, or information that was "known" but not recorded, could not be represented in the term lists. Clinician-maintained problem lists and patient medication information would lead to richer term lists. Better filtering of MeSH terms could eliminate non-useful terms.

Even with more complete extraction of information and better mapping to MeSH terms, the information retrieved will be limited by the way underlying knowledge sources are indexed and retrieved. One could create a complex model of a patient that takes into account all co-existing diseases, risk factors, and medical and surgical treatments. But with current indexing methods, whether using keywords, MeSH terms, or full-text indexing, the patient model must be converted back into a group of words or terms, with or without a Boolean construct, and thus loses the richness inherent in the clinical record. A new way of indexing medical knowledge sources, one that would permit a richer retrieval system based on a model of a patient's current medical circumstance, might enhance the usefulness of information retrieval applications connected to the EMR. Incorporation of richer clinical terminologies designed to represent patient-specific information, such as SNOMED, might yield better search results, especially in resources that have not been indexed using MeSH.

Our work builds upon, diverges from, and extends previous work in several ways. SmartQuery somewhat resembles CHARTLINE, which used the words in a textual patient record to map to UMLS terms containing those words, then used co-occurrence data and heuristics to generate pairs of co-occurring terms that might provide useful MEDLINE searches.[15] We also use text in the record, though only from selected sections of the most recent admission H&P, to identify MeSH terms of interest. We extend that work by allowing the user to combine terms accumulated from disparate parts of the EMR,

and by providing simultaneous queries to a variety of resources in addition to MEDLINE.

Our system also resembles the MEDLINE button and subsequent Infobuttons developed by Cimino and colleagues in that it uses a list of ICD-9 codes as a problem list and attempts to map those codes to MeSH terms.[8, 16, 17] Unlike their system, we do not use generic queries. We try to elicit terms that describe the clinical situation, then help the user query various resources. We hypothesize that by allowing users to accumulate information from multiple areas of the EMR, and to contribute their own search terms, we can facilitate a richer collection of terms that can be combined to produce queries involving multiple concepts. We also allow the user to select which resource(s) he would like to search, allowing for a more individualized search. In the absence of formal evaluation data, it is difficult to know which of these approaches best meets the combined needs for efficiency and successful retrieval.

Our evaluation was preliminary and had limitations. The study was limited to a specific group of users and employed simulated clinical tasks. SmartQuery's usefulness in other specialty domains, other clinician groups, and real clinical environments has not been tested, and we did not compare task performance with and without SmartQuery. The residents who volunteered to participate in the study may have been favorably biased and may not be representative of a more general population. Furthermore, the opinions expressed on the questionnaire may not reflect what usage would be in daily clinical situations. Further research should also test the assumption that making knowledge resources conveniently accessible will increase their use, or that use will lead to improved patient outcomes or to more cost-effective care.

CONCLUSIONS

Our work suggests that provision of context-sensitive links from the EMR to relevant knowledge resources can improve access to information in the clinical setting. We have shown, as have others, that it is feasible to predict some of the topics that will be of interest to clinicians caring for a particular patient, and to provide direct links to information about those topics and about topics entered directly by the user. Further work is needed to improve search term prediction and query formulation.

REFERENCES

1. Gorman P and Helfand M. Information seeking in primary care: How physicians choose which clinical questions to pursue and which to leave unanswered. Med Decis Making 1995;15:113-9.
2. Covell DG, Uman GC and Manning PR. Information needs in office practice: Are they being met? Ann Intern Med 1985;103:596-9.
3. Ely JW, Osheroff JA, Ebell MH, et al. Analysis of questions asked by family doctors regarding patient care. BMJ 1999;319:358-61.
4. Hersh WR and Hickam DH. How well do physicians use electronic information retrieval systems? JAMA 1998;280(15):1347-52.
5. Ely JW, Levy BT and Hartz A. What clinical information resources are available in family physicians' offices. J Fam Pract 1999;48(2):135-9.
6. Osheroff JA and Bankowitz RA. Physicians' use of computer software in answering clinical questions. Bull Med Libr Assoc 1993;81(1):11-9.
7. Gorman PN, Ash J and Wykoff L. Can primary care physicians' questions be answered using the medical journal literature? Bull Med Libr Assoc 1994;82(2):140-6.
8. Cimino JJ. Linking patient information systems to bibliographic resources. Meth Inform Med 1996;35:122-6.
9. Hersh WR and Donohoe LC. SAPHIRE International: A tool for cross-language information retrieval. Proc AMIA Annu Fall Symp 1998:673-7.
10. Braunwald E, Fauci AS, Isselbacher KI, et al., editors. Harrison's Online: McGraw-Hill; 2001. http://www.harrisonsonline.com
11. National Guidelines Clearinghouse. Accessed: August 9, 2001. http://www.guidelines.gov
12. Hersh WR, Brown KE, Donohoe LC, Campbell EM and Horacek AE. CliniWeb: Managing clinical information on the World Wide Web. J Am Med Inform Assoc 1996;3(4):273-80. http://www.ohsu.edu/cliniweb
13. Payne TH, Murphy GR and Salazar AA. How well does ICD9 represent phrases used in the medical record problem list? 16th Annu Symp Comput Appl Med Care 1992:654-7.
14. Cimino JJ and Sideli RV. Using the UMLS to bring the library to the bedside. Med Decis Making 1991;11 (suppl):S116-20.
15. Miller RA, Gieszczykiewicz, Vries JK and Cooper GF. CHARTLINE: Providing bibliographic references relevant to patient charts using the UMLS metathesaurus knowledge sources. 16th Annu Symp Comput Appl Med Care 1992:86-90.
16. Cimino JJ, Johnson SB, Aguirre A, Roderer N and Clayton PD. The Medline button. Proc 16th Annu Symp Comput Appl Med Care 1992:81-5.
17. Cimino JJ, Elhanan G and Zeng Q. Supporting infobuttons with terminological knowledge. Proc AMIA Annu Fall Symp 1997:528-32.