Richard P. Medlin. Providing better clinical reference sources at the point of care. A Master's Paper for the M.S.I.S. degree. November 2013. 60 pages. Advisor: Javed Mostafa

Physicians frequently need clinical reference material at the time they are taking care of patients, but despite the widespread availability of seemingly appropriate resources, physicians do not consistently avail themselves. Previous research has shown that the main obstacle to resource use is lack of time to search, followed by inability to find the information even with extended searching.

Using a randomly selected sample of previously published questions (Clinical Questions Collection at the National Library of Medicine), current online electronic clinical reference resources were searched. Search times, success rates, and result quality were compared. Retrieval of drug information was dominated by navigation time, whereas searches for other materials were dominated by search time. It was rare that actually reading the material required significant time.

Based on the results, an improved interface for accessing clinical reference data is proposed. Key features include a "Fast" interface that minimizes navigation time to easily found resources and a "Flex" interface that minimizes search time for more arcane material.

Headings:

Medical informatics

Information-seeking behavior

Information needs

Electronic health records

Decision support systems

Electronic information resource searching

PROVIDING BETTER CLINICAL REFERENCE SOURCES AT THE POINT OF CARE

by
Richard P. Medlin Jr.

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

December 2013

Approved by

_____

Javed Mostafa

Table of Contents

Introduction

When selecting therapy for patients, physicians frequently have unanswered questions regarding the best course of medical care. This issue has been studied repeatedly over the last 25 years, but despite the increasing availability of robust electronic reference sources, the number of unanswered questions remains alarmingly high. In Covell's landmark study in 1985 (prior to the widespread availability of electronic medical reference sources), he found that family physicians had unanswered questions during approximately 66% of patient encounters in a primary care setting (Covell, Uman, & Manning, 1985). Covell had originally speculated that the problem was related to inadequate reference material, and called for the creation of electronic reference sources. However, data collected in 2004, after the advent of widespread availability of electronic textbooks (accessed either on local DVD or via the internet), showed a similar number of unanswered questions (Ely, 2004; Graber, Randles, Ely, & Monnahan, 2008). In both Ely's and Graber et al.'s studies, numerous electronic resources were available; however, physicians made a conscious decision not to pursue the answer to a question through the available resources in roughly 50% of cases. The major reason cited was a lack of time, but other contributing reasons were doubt that a published answer to the question existed, and that physicians did not think that the answer to the question would significantly impact the care of that specific patient. A recent grounded theory-based study found

similar reasons for not pursuing answers to questions regarding patient care (Cook, Sorensen, Wilkinson, & Berger, 2013).

In addition to not pursuing answers to clinical questions, physicians have also been overly confident regarding their platform of knowledge. A study by Meyer had physicians review a series of case vignettes and select a diagnosis. The cases increased in difficulty, with 58% of the physicians making the correct diagnosis for the easier cases, but only 6% making the correct diagnosis for the more difficult cases. Despite the staggering drop in diagnostic accuracy, physicians' confidence in their diagnoses remained high (Figure 1) (Meyer, Payne, Meeks, Rao, & Singh, 2013). This suggests that physicians have significant knowledge gaps of which they are unaware ("unknown unknowns.")

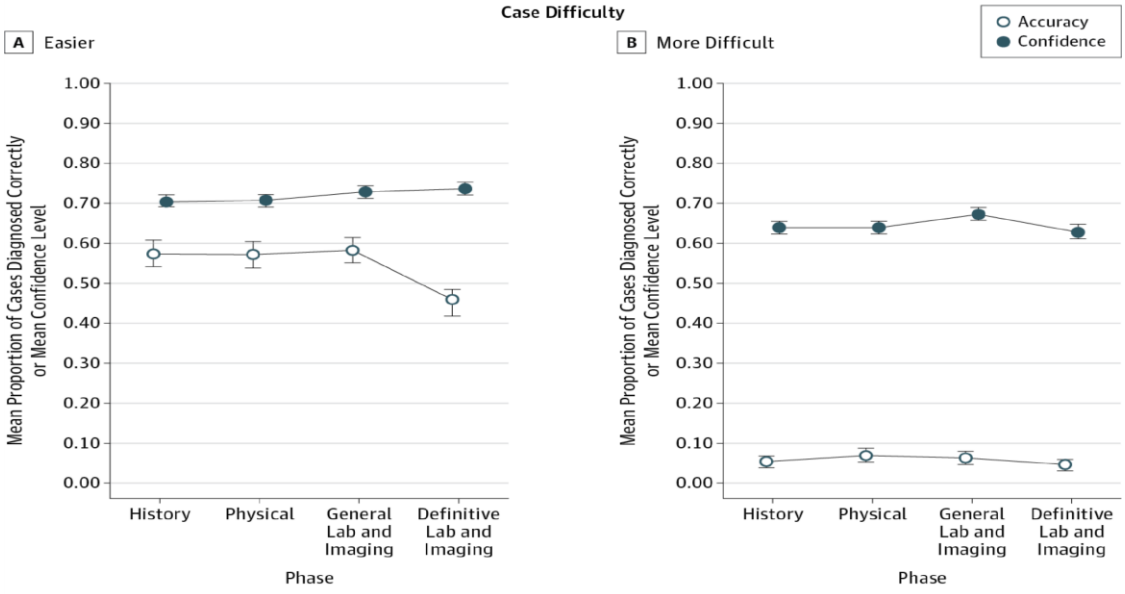


Figure 1 - Physicians' mean diagnostic accuracy and confidence in that accuracy as a function of diagnostic phase and case difficulty (easier vs. more difficult). Lab indicates laboratory testing; error bars represent ±1 SEM (from Meyer et al., 2013).

Even when physicians do pursue a question, their searches may lead them to the wrong answer. A laboratory study of Emergency Medicine residents demonstrated that when

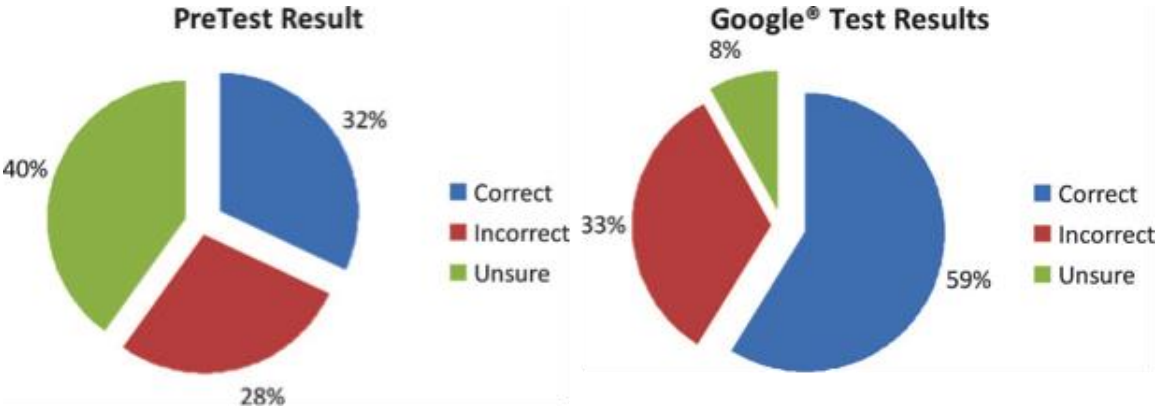


Figure 2- Google quickly satisfies information needs. Unfortunately, users are over-confident regarding the accuracy of the information retrieved (Krause et al., 2011).

pursuing answers to questions using only Google® (as opposed to directly searching clinically oriented and authoritative references, such as UpToDate®), users usually found an answer that they felt comfortable enough to act on 90% of the time. Unfortunately, the answer was incorrect in 30% of cases. The users who found the "wrong" answer generally transitioned from "I don't know the answer" to "I am confident about my [incorrect] answer" (Abbas, Schwartz, & Krause, 2010; Krause, Moscati, & Halpern, 2011).

When reviewing the search strategies from this study, the authors noted that the Emergency Medicine residents frequently referenced consumer-level websites (e.g., MedlinePlus, wrongdiagnosis.com), never used advanced Google search syntax, and never proceeded to the university library when their search ended up at a paywalled content provider like ScienceDirect (Abbas et al., 2010).

In summary, lack of time has continued to be a problem for clinician reference resources at the point of care with regard to physician information-seeking behaviors. The addition of electronic resources as a supplement to or in lieu of the traditional printed material has resulted in physicians looking for more answers, but often locating and applying incorrect answers, which likely results in suboptimal patient care. Physicians will use general internet search engines such as Google even when professional level references are available, presumably because Google provides affordances not currently offered by clinically oriented references. Finally, physicians frequently have unmet information needs of which they are unaware, which suggests that some information may need to be "pushed" to physicians, rather than depending on physicians to institute a search.

In addition to the above-noted problems with physician retrieval strategies, the amount of accessible clinical reference information that is both relevant and retrievable has increased. Even the most dedicated physician cannot keep up with all of the research published. The traditional approach to this data overload has been that physicians become specialists or subspecialists, and thus limit the knowledge domain for which they are responsible. This approach appears to be effective in actually improving outcomes for some conditions (Cobin, 2002; Eaton, Murphy, & Hunt, 1997; Goldstein, Matchar, Hoff-Lindquist, Samsa, & Horner, 2003). However, it has practical limitations in terms of how specialized a physician can become (Nash, Josephson, Sun, & Ferriero, 2013) and how accessible highly specialized physicians are to the general patient population. In addition, the rate of creation of new subspecialties could not possibly keep up with the rate of knowledge generation through medical research.

While the pace of knowledge generation is impressive, the rate of data accumulation

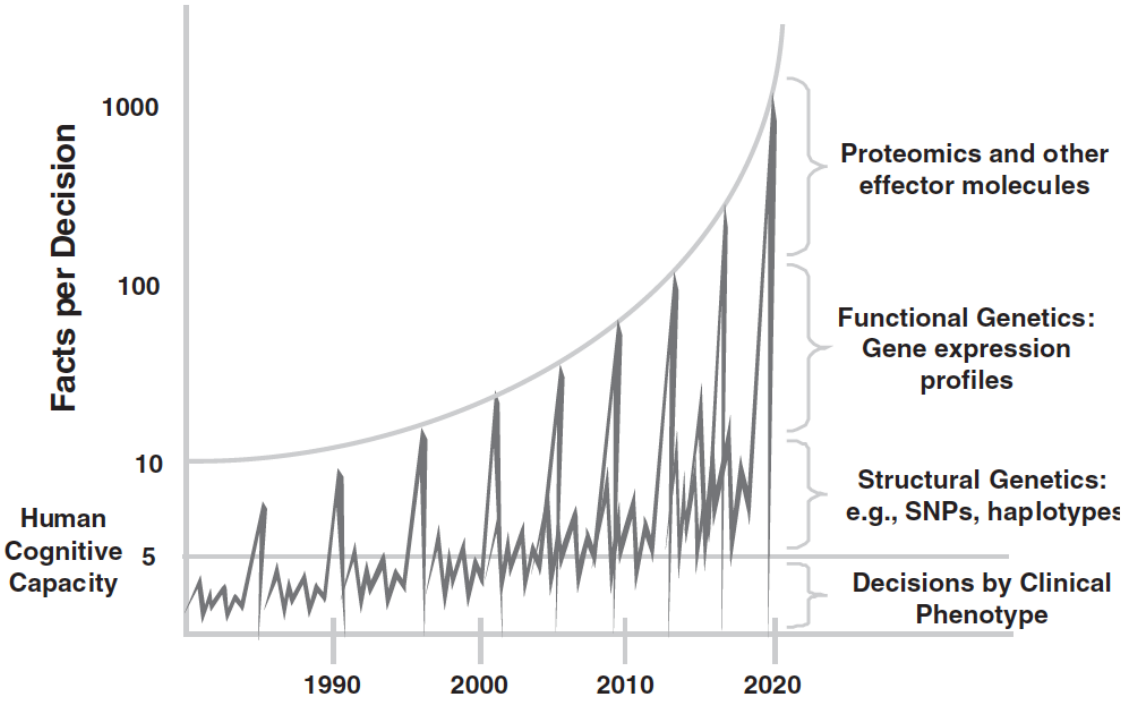


Figure 3 - Schematic depicting the increase in number of facts per clinical decision with new sources of biological data. (McClellan et al., 2008, p.19)

within the medical records of individual patients is accelerating even more rapidly. The first big jump in data quantity occurred with the implementation of electronic health records. Physicians now routinely have access to nearly complete medical records for many patients, including all hospital and clinic data, which was rarely the case ten years ago. In addition to the written records generated by clinical staff, machine-generated data in the form of laboratory results, diagnostic images, and medical device data logs from devices such as pacemakers and insulin pumps are included in the medical record. Ultimately, the addition of genetic sequence data to the patient's medical record will amplify the number of facts that need to be considered for each clinical decision in an exponential fashion (Figure 3), which will outstrip human cognitive capacity. The best

medical decisions, which consider all of the possible interacting factors, will only be able to be made with the assistance of a machine. Arguably, this is happening already. Sophisticated image processing software is used to produce readable images for radiologists. The output of these imaging systems is hardly a reflection of the "raw data" produced by an MRI scanner.

Herbert Simon noted that in an information-rich environment, human attention and processing becomes the limiting factor in the performance of any task. At some point, aggregating and presenting additional information will only lead to worsening performance. Once that point is reached, the only useful information-processing systems are ones that actually limit the amount of information that reaches the user (Simon,1971). An example of this effect occurred in the airline industry during the production of the Boeing 777 airliner. After considerable study, it was determined that adding more gauges and controls in the cockpit actually degraded pilot performance. The resulting design of the cockpit was an instrument package that did not reveal all data available on the "home screen" and "fly-by-wire" controls that allowed considerable automation to be incorporated into the actual operation of the flight control surfaces (Abbott, 2001). Similarly, physicians have now reached the point at which additional information is likely to result in worsening performance. It is clear that the implementation of robust Clinical Decision Support (CDS) systems will be required for preventing "data overload" in clinical environments (McClellan, McGinnis, Nabel, Olsen, 2008). Careful design will be required to optimize what data is filtered and presented. Appropriate automation of certain tasks is likely to result in improved physician decision making.

The general term Clinical Decision Support (CDS) is used to describe a broad array of resources and techniques that provide patient-specific guidance to clinicians at the point of care. In the past, "low tech" solutions such as printed clinical guidelines and pocket handbooks were used. The surfeit of new information (as depicted in Figure 3), including both medical knowledge and patient-specific facts, has made the implementation of computerized decision support systems a requirement to effect complete and appropriate care decisions in the medical setting. The vision of a technology-enabled Clinical Decision Support system has been described by Centers for Medicare and Medicaid Services as:

> "[CDS] requires computable biomedical knowledge, person-specific data, and a reasoning or inferencing mechanism that combines knowledge and data to generate and present helpful information to clinicians as care is being delivered. This information must be filtered, organized, and presented in a way that supports the current workflow, allowing the user to make an informed decision quickly and take action. Different types of CDS may be ideal for different processes of care in different settings." (www.healthit.gov, 2013)

The ability to implement CDS systems is predicated on data being available in a form that is usable by computers. While much of the recent patient data (particularly quantitative data, such as lab results) was "born digital" and thus is readily consumed by CDS systems, most medical knowledge and a great deal of individual patient data is not easily converted to a computable form. Even when stored electronically, much of the information is unstructured free text, tables, or images, all of which present considerable challenges when preparing them for use by reasoning systems (including human reasoning systems.) Therefore, incremental attempts have been made to implement CDS systems, either by focusing on the parts of the medical record and medical knowledge

base that are computable (e.g., medications, labs, and genetics) or by delivering less-structured results and letting the clinician provide a more flexible reasoning system. The focus of this paper is on the latter type of system.

Literature Review

It has been quite difficult to implement effective Clinical Decision Support systems for the vast majority of patient care. First, relative to the number of diseases that occur, there are relatively few diseases which have any evidence-based treatment recommendations

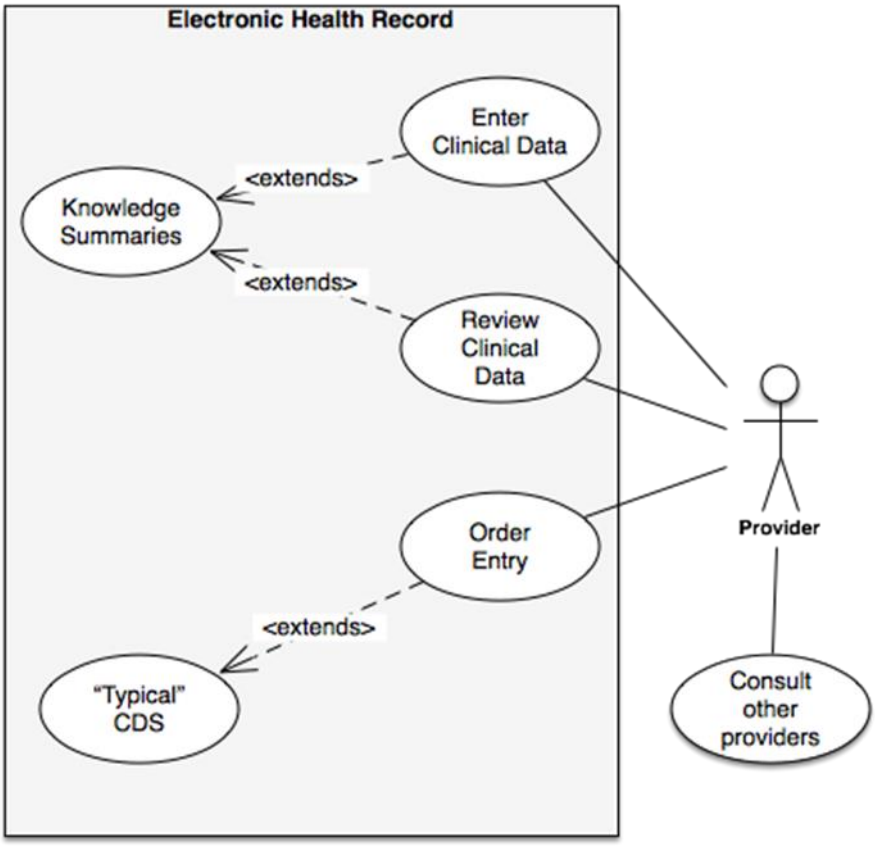


Figure 4 - Clinical decision support use case – There are many methods for implementing CDS. Many of them are implemented by leveraging the clinician's interactions with the EHR. In addition to the usual CPOE-based CDS, one could implement "Knowledge Summaries" containing specific information germane to that particular patient. Computerized CDS always competes with the traditional CDS of consulting a colleague. The higher the threshold for consults, the more likely physicians are to use other reference sources.

other than expert opinion (McClellan et al., 2008). Second, most of the guidelines that have been produced are not in a form that is easily interpreted by the physician, much less a computer programmer (e.g., the phrase "consider consulting surgery" is neither an actionable recommendation nor does it translate well to computer code). Last of all, most patient data is inherently in unstructured, narrative form that cannot be utilized directly by computer algorithms to drive selection of appropriate guidelines ("Mrs. Jones states that she feels better than her last visit" requires a considerable number of contextual cues to interpret correctly.) For the foreseeable future, it is likely that a "human in the loop" will be needed to make sense of patient stories. Currently implemented CDS available in most commercial EHRs include alerts, reminders, order sets, drug-dose calculations that automatically remind the clinician of a specific action, or care-summary dashboards that provide performance feedback on quality indicators Links to clinical information sources, such as hyperlinks to the local medical library or to a specific guideline associated with an order set, are also quite common. Prescribing databases (Lexicomp, Epocrates, or Micromedix) are also generally available. However, all of these information sources require a clinician to both formulate and execute a query, should they decide that they need information.

The most successful instances of CDS to date involve measures that do not require active initiation by the physician, which have a comprehensive set of knowledge in structured format, and which have some rules that are both easily expressed in computer logic and where the recommendations have a high probability of being relevant. Additionally, in order to show that CDS is working, the outcome variable needs to be easily measurable.

The one area in which all of these criteria are met is medication management. Drug selection and prescriptions have been obvious targets for CDS not only because drug information is already structured and adequately detailed, but also because the placement of orders for medications occur in a predictable and easily identified point in the clinical workflow and it has recently become obligatory for physicians to prescribe electronically (Computerized Provider Order Entry - CPOE), rather than writing the prescription or order on paper. This means that the physician will be present and (presumably) attentive to the computer during this process, which affords an opportunity to give feedback directly to the provider at the time of the prescription.

While the "best" drug for a given condition may be hard to choose, it is considerably more straightforward to detect "bad" drug choices – e.g., a drug to which the patient is allergic, the wrong dose for the patient weight, wrong dosing interval, or interactions with other medications. Detecting these sorts of errors has been the primary focus medication CDS, rather than making sure that the physician is choosing the "best" drug. The combination of CPOE and CDS has resulted in a substantial decrease in the number of potentially harmful medication prescribing errors. It is important to note that one of the reasons that this intervention has been deemed successful is that it is easy to measure and quantify medication errors. However, there have been new errors introduced by CPOE (dubbed "e-Iatrogenesis" by Weiner et al.). Virtually all physicians who have transitioned from paper prescribing to electronic prescribing have had the experience of erroneously choosing the wrong medication from a long drop-down list, resulting in

selection of the wrong medication or dosage. This is an error that simply cannot occur when writing out the name of the drug by hand (Bright et al., 2012).

The other problem that has occurred with medication-related CDS is that, despite the seemingly straightforward rules that need to be implemented, there is still an unacceptably high incidence of false alarms. Each of these false-positive alerts requires considerable interaction from the physician to overcome, typically by typing a reason for overriding the alert in a dialog box. "Alert fatigue" occurs from repeated presentation of irrelevant or frankly incorrect information during computerized order entry, leading clinicians to not only ignore the alerts but to even request that they be removed because they impede workflow. Even the best systems can fail in the face of some specific scenarios: the seemingly straightforward task of alerting a physician to a patient's allergy can fail for a specific patient due to the subtleties of the specific clinical situation. For example, an order for cefazolin in a patient allergic to penicillin will inevitably trigger an incorrect alert because the system does not account for (1) the reliability of the patient's report of the penicillin allergy[1], (2) the low probability of cross-reaction between cefazolin and penicillin (less than 5%), (3) the seriousness of the current patient's illness (e.g., dying of sepsis), and (4) the lack of alternative treatments. While this example may seem like it requires a confluence of improbable events, it is certainly one of the more common false alerts experienced by the author.

[1] Patients who report previous allergies to penicillin rarely have actual allergies when evaluated with skin testing.

In the final analysis, there are very few decisions in medical practice that have clearly computable answers, either because the inputs are unstructured or because there is no appropriate evidence-based algorithm to apply. In fact, there have been studies that suggest that when expert clinicians are forced to apply clinical guidelines in a strict fashion, their patient outcomes revert to the level of a novice clinician, which is clearly not the desired outcome (Smith, 2003). In other words, in order to provide useful guidance and cognitive support to physicians, it is necessary to know a great deal of supporting information about the patient, the care context, the clinician, and perhaps to be able to incorporate tacit knowledge that is not available in any reference material.

In order to sidestep the need for extensive structured data, less-proscriptive clinical decision support, such as simply supplying links to clinical reference information in the EHR interface, has been implemented frequently. Prior to the widespread availability of the Internet, physicians universally carried reference books that were sized to fit in a lab coat pocket. The downside of these types of resources, in both physical and electronic form, is that the physician needs to know that they have an information need, and they need to believe that they have both the time and the correct resource to answer the question at hand. As noted above, physicians are frequently unaware that they do not possess critical knowledge needed for patient care (Dhaliwal, 2013), and even when they are aware of their knowledge deficit, they may not feel that they will be able to find it in a reasonable amount of time (Ely et al., 2000 ; Cook et al., 2013; Jerome et al., 2001).

Various techniques have been used to encourage physicians to pursue their questions. Many systems focus on making searching less time-consuming. For example, it makes sense to link directly to drug dosing information during the prescribing activity in an EHR or to link to a lab reference during the lab review activity. In fact, these were two of the most popular uses of electronic reference materials in one study (Chen & Cimino, 2003).

Another way to speed up searches is to predict and perform searches of reference materials that are likely to occur, without requiring the physician to construct a query. Some early systems tried to predict queries based on the content of the patient medical record. One of the earliest systems automatically filled in generic MEDLINE queries with specific information from the patient's medical record (Cimino et al., 1992). For example, the query "what is the treatment for disease X?" would substitute diseases from the patient's diagnosis list for "X". Thus, pre-built queries for a patient with diabetes and hypertension would include "What is the treatment for diabetes?" and "What is the treatment for hypertension?" These could be selected by the physician instead of forcing the physician to construct the complicated query necessitated by early MEDLINE search software. This system was used with real patient data, but never put into production.

Another attempt at automatically generating queries was based on indexing the content of the free-text portion of medical records using TF*IDF. The index terms generated for each patient's medical record were used to "fill in" generic queries of the type in the study above. These were compared to keywords that physicians developed by manually

reviewing the records to fill in the same queries. They found a significant correlation between the documents retrieved by the clinician-generated terms and those retrieved by the automatically indexed terms, but the overall agreement was rather low, ranging from 0.04 to 0.52, depending on the question (Mendonça, Cimino, & Johnson, 2002).

A more recent experimental system used other structured data from the medical record to generate queries (age, gender, abnormal labs, and current active diagnoses). These queries were incorporated into a search interface that was designed to retrieve documents from PubMed that could be used in an evidence-based medicine analysis of a clinical question. The underlying algorithm for re-indexing PubMed was essentially TF*IDF (BM25). During development, they found that including numerous very specific search terms tended to result, unsurprisingly, in empty retrieval sets, so they added "fuzzy matching" logic to some of the query terms in order to improve performance (Krumpholz, 2012). The system was not evaluated quantitatively, but a focus group of physicians felt that it improved their search experience.

One of the few actual production systems that provides access to clinical reference material and utilizes contextual information from the EHR is the Infobutton system, developed by Cimino et al. Infobuttons have been implemented at the VA Medical Center, Columbia University, and the University of Utah. Several commercial vendors have developed interfaces, and some of them are in production. Most recently, an Infobutton standard has been developed and approved by HL7 (an important Health IT

standards organization). This standard allows an EHR to simultaneously query multiple commercial and publicly available knowledge sources via an HL7 compliant interface.

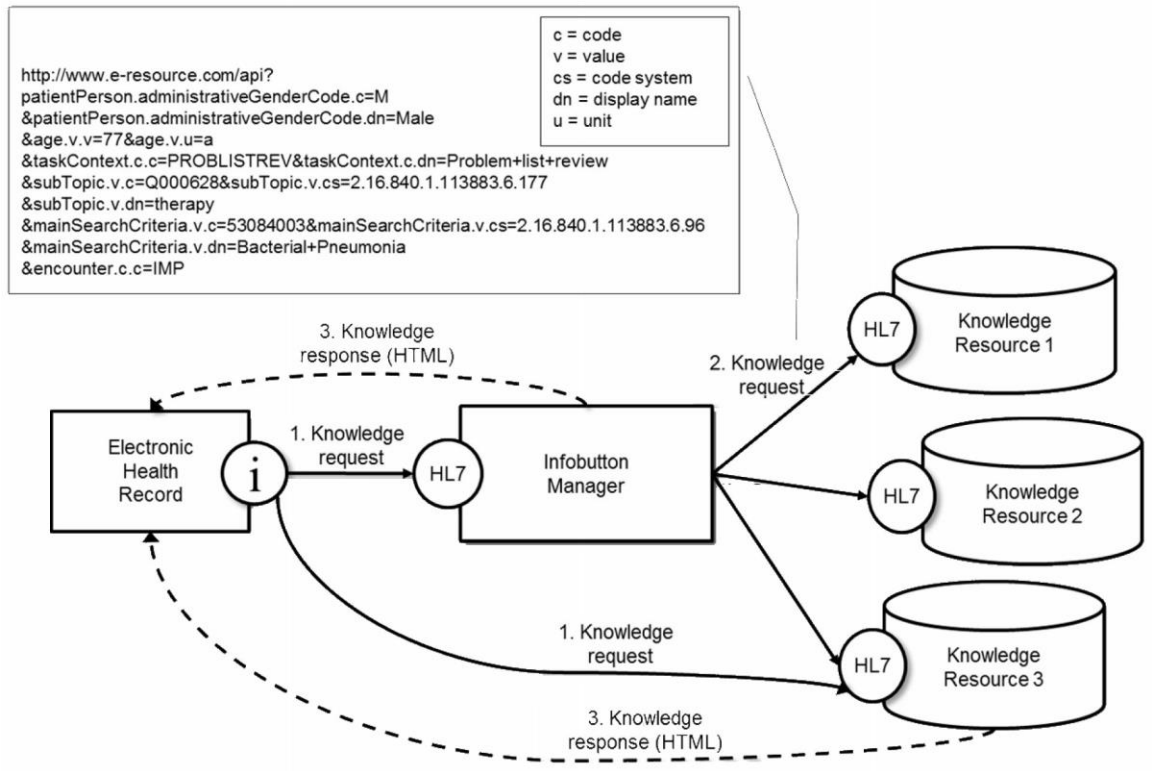


Figure 5 - The OpenInfoButton webservice architecture. Note that the message to the knowledge resources contains information that can be derived from the EHR context including (1) the patient (e.g., gender, age); (2) the clinical information system user (e.g., discipline, specialty, preferred language); (3) the task being carried out in the clinical information system (e.g., order en- try, problem list review, laboratory test result review); (4) the care setting (e.g., outpatient, inpatient, intensive care); and (5) the clinical concept of interest (e.g., a medication order, a laboratory test result, a problem). (from Del Fiol et al, 2012)

From the user perspective, Infobuttons provide a uniform search interface experience across different organizations, different EHRs, and different knowledge sources, which decreases the learning curve for providers. Infobuttons also only reference authoritative sources, so it is less likely that these resources will result in the degree of erroneous

information retrieval that was found when physicians used general purpose internet search engines. Although this has not been described, one outcome variable that could be measured after implementing an Infobutton system would be a decrease in the use of Google for clinical reference queries.

The Infobutton standard does have some limitations, including the inability to submit queries with sophisticated logic and the lack of control over how the knowledge sources index their content. It is likely that different knowledge sources will return wildly different results for ambiguous or obscure queries, due to indexing differences.

The downside of all of the systems noted above, including Infobuttons, is that physicians still need to know that they have an information need and they need to believe that they have both the time and the correct resource at hand to answer the question. As noted above, physicians are frequently unaware that they do not possess critical knowledge needed for patient care (Dhaliwal, 2013), and even when they are aware of their knowledge deficit, they may not feel that they will be able to find it in a reasonable amount of time (Ely et al., 2000; Cook et al., 2013; Jerome et al., 2001). Therefore, the next logical phase in clinical reference retrieval is to retrieve and present information that users are not yet aware that they need, thus converting the clinical reference activity from a "pull" activity to a "push" activity.

Guilherme Del Fiol, one of the developers of the Infobutton standard, was interested in designing and prototyping a search interface that would extend what could be delivered

to providers via the Infobutton standard. In particular, he wanted to deliver highly condensed "Knowledge Summaries," combining intelligent summarization and the merging of different types of clinical reference content, that were tailored based on the contextual elements derived from the EHR. The following research is a pilot project that investigated the ability to produce these summaries using manually curated materials, and then a preliminary investigation into how to present this information to users.

Methods (Pilot Study)

We performed a pilot study to determine if the questions that Ely had originally classified as "difficult to answer" or "unanswerable with online resources" were still difficult or unanswerable. The questions that Ely originally collected are available via the National Library of Medicine's Clinical Question Bank (http://clinques.nlm.nih.gov), along with rich metadata regarding the practice, the patient, and the physician who asked each question. We selected 50 questions at random from the question bank, and I attempted to answer them using a convenience sample of current commercial clinical reference resources (UpToDate™, DynaMed®), a government-sponsored clearinghouse (Guideline.gov), PubMed, a commercial medication database (Lexicomp™), and a general purpose internet search engine (Google). Three of these were resources which I had not used prior to this study (Dynamed, Guideline.gov, and Lexicomp).

My hypothesis was that nearly all of the questions would now be answerable because (1) questions regarding then-cutting-edge clinical concerns were likely to have been definitively resolved, e.g., "Does Prozac cause cancer?"; (2) more comprehensive content is now available; (3) authors have started including more clinically actionable content in their writing; (4) the information retrieval and indexing algorithms have improved; (5) user search skills have improved.

Overall, I believed that commercial sources (UpToDate, DynaMed) would not only have at least 90% of the answers, but also that these answers would be easy to find due to extensive internal search engine optimization and internal hyperlinking of content. I included a specific drug reference (Lexicomp) in the resource list which I believed would answer all drug-specific questions (dose, side effects, indications, interactions), but I considered it likely that I could find all the information in Lexicomp via Google (from the manufacturer, the FDA, Drugs.com, WebMD, and other curated collections of medication data), and that Lexicomp's main advantage would be the consistent structure of the drug monographs, which would allow for more rapid reading.

I attempted to use each resource to answer each question. The sequence of resources I used was randomized for each question. The outcome of a search could be 1 of 5 choices: "Complete Answer," "Partial Answer," "No Answer," "Wrong/Conflicting Answer," or "Serendipitous Answer." An example of a serendipitous answer was when I found a better answer than would have been available if I had strictly answered the question. This typically occurred when I was looking up a drug dose and found that there was a better drug for the indication.

I recorded the following data for each question:

- The text and ID of the question
- My *a priori* answer (answering from my own knowledge)

- My *a priori* outcome (my answer after completing all searches; "complete answer," "partial answer," "no answer," "conflicting/wrong answer," "serendipitous answer")
- Given both the question and the answer I found, would I send this patient to a specialist?
- For each search on each resource I recorded:
    - Outcome ("complete answer," "partial answer," "no answer," "conflicting/wrong answer," "serendipitous answer")
    - Was the search abandoned without finding an answer?
    - How much time was spent on the search
    - URL of the final answer
    - Search terms used
    - Content fragment that contained the answer
    - Section heading, if applicable
    - Authoritative source of final answer, if able to determine (e.g., original literature cited? If so, what was the PMID?)
    - Any comments about the search process.

I used Morae (TechSmith, Inc, Okemos, MI, www.techsmith.com) to record the amount of time that was spent searching, navigating, and then reading each document. When I found an answer, I used the cursor to select the part of the document that answered the question. If multiple documents were needed, I did not record the question as answered

until I had reached the last document. The total time for each search was recorded from the time I finished reading the question to the time of the last cursor selection.

I generally searched for 10 minutes and declared the search "abandoned" if I did not seem to be making any progress. In the case of Lexicomp, which is primarily a medication database, I abandoned the search with less than 10 minutes of searching if the initial search did not return any information and the domain of the question was not related to medication administration. I also recorded extensive notes regarding my search strategies and noted when my searches had unexpected results.

Results (Pilot Study)

I answered 25 questions using all 6 resources. My *a priori* knowledge of the questions allowed me to completely answer 4 questions, partially answer 11 questions, and incorrectly answer 1 question. I could not answer the remaining 9 questions. I am sure that my *a priori* knowledge likely affected my search strategies, particularly with PubMed and Google. The searches lasted between 20 seconds and 32 minutes (researching a basic science question). Sometimes, the searches ended with a copy of a very similar document being retrieved from different repositories, particularly for Google and PubMed (e.g., an author's proof vs. the publisher's final document.

Identical content was present in several resources

For "simple" medication queries (e.g., dose, route, interactions, side effects), I found that all 3 of the commercial resources were backed by the same data source (American Hospital Formulary Society's Drug Information monographs - AHFS DI®), although each resource had some minor variations in presentation and formatting. DynaMed and Lexicomp directly accessed the AHFS resource; UpToDate indirectly linked to AHFS DI by linking to Lexicomp, which then displayed data from the original AHFS datasource. Therefore, all 3 commercial vendors used essentially the same medication resource.

I found it difficult to find medication information in which I was confident by using Google. Many of the answers I found via Google were actually correct, but I would not have adequate confidence to act on them in a clinical situation. The FDA-approved Prescribing Information (PI) was surprisingly difficult to find, probably because many of the drugs which I was looking up were generic. I did not come across any free comprehensive sources of off-label dosing information in which I felt confident. Overall, the only resource that had this information was AHFS DI.

In addition to duplicate medication content, other sources frequently cited or linked to information taken from the Cochrane Library meta-analysis, professional society practice guidelines (e.g., Advanced Cardiac Life Support), or disease-specific society guidelines (e.g., the American Diabetes Association). If there was a seminal article on a particular subject, DynaMed, UpToDate, and Guideline.gov usually linked to the PubMed reference, but not to the original article, even if it was freely available on the author's website.

Google Scholar indexes all of the PubMed content, but ranks documents differently

Google also occasionally linked directly to original research from the main search page, but it usually linked to the Google Scholar search interface, where the search was re-run with the same search terms. I was frequently able to find articles in PubMed via the Google Scholar search engine that I was unable to find via PubMed's own search engine. Additionally, Google Scholar frequently linked directly to a free copy of the article (although the legality of some of these links was suspect), whereas PubMed linked to a

paywalled publisher's website. Additionally, I inferred that publishers must supply full-text content to Google to index, because Google would sometimes display snippets in the search results that were from behind a paywall.

On simple queries, navigation takes much longer than assimilating the answer

One interesting finding was that the majority of the time spent searching for information on simple medication queries (e.g., what is the adult dose of amoxicillin for sinusitis?) was actually spent navigating to, authenticating, and then navigating through the results, not actually reviewing material; the navigation sometimes took 2 to 3 times longer than reviewing the information. This observation had a significant impact on the design of the knowledge summary interface.

There are more electronic resources available in 2013 than in 2004. The content contains more actionable information but there is still no resource with "all the answers"

In the original Ely study, 20% of the questions could not be answered by any resource, and the best resource only had 7% of the answers. In this pilot study, all questions could be answered by a combination of resources, and the best resource had 72% of the answers (UpToDate). However, only one question was answered by all 6 resources ("What is the best treatment for bed-wetting?"), and there was one question which had an answer in only one resource ("Why does the same virus cause a cough in a kid and just rhinorrhea in an adult?")

20% of the questions required reviewing multiple resources or articles to get a complete answer to the question. In 20% of the cases, the answer to the question would have included referral to a specialist. This is similar to the findings of an e-mail consultation service which found that only 12% of the questions resulted in a referral and 7% were deemed "unanswerable" (Bergus, Emerson, Reed, & Attaluri, 2006).

The relevant information was often contained in a table or image

Critical information was frequently contained in non-textual content. Medication recommendations were frequently presented in a table (which was saved as an image). Treatment guidelines were often represented with an annotated flowchart. The results of primary studies were often best digested as a graph (particularly a Kaplan-Meier survival curve), and meta-analysis results were often presented as a forest chart. While it is possible to extract text from tables or images computationally (for indexing purposes), it does not seem useful to present only the indexed text to the user, as the user is unlikely to find that format usable. Whatever the final interface, the ability to display graphic images seems critical.

Document level retrieval was inadequately granular for efficient information assimilation

The primary documents retrieved were sometimes quite long (over 100 pages for some guidelines), and this resulted in significant navigation within documents, particularly those retrieved from Google. Even in UpToDate, DynaMed, and Lexicomp, which

should be optimized for clinical use, I had to pursue several links to find the desired content, even if I was on the correct page. When using Google, I sometimes could not find the phrase I saw in the Google-supplied snippet in the referenced document because the text had been indexed from an image, table, or some other metadata that was not easily searchable using the browser-based “search this page” function.

User Interface Design

Based on the results of the pilot study, we concluded that the answers for most questions existed online, although there was no single comprehensive resource that had all of the answers. The initial purpose of the project was to develop a user interface to display the knowledge summaries. In light of the additional insight gained during the pilot study, we added additional components to the initial design goal of "display important reference information in a compact form." For the purposes of designing the interface, we made several assumptions

- We would eventually be able to retrieve content using computational means that was comparable with what we produced with manual retrieval.
- We could pick out portions of the document that were most relevant to the current clinical situation and that our relevance judgments would be broadly applicable for other similar clinicians.
- The information displayed was situationally relevant to the user (Pluye et al., 2005).
- We had access to both commercial and public content.
- We assumed that we had access to the patient's chart and some information about the physician and setting (e.g., primary care physician, more than 10 years out of residency, in the context of an office visit for an established patient.)

We felt that there are certain questions, based both on published studies and our own experience, that are best answered during the process of order entry (for example, "What is the default dose for this patient's age and weight?" is "answered" by simply having the computer supply the correct default dose.) For other medication-related questions, however, it would be important to have the information available before starting the prescribing dialog. For instance, physicians frequently want a list of alternative medications for the same diagnosis because the patient may have a comorbid illness that might influence medication selection. For example, a patient who has migraines and is hypertensive might benefit from being on verapamil, an anti-hypertensive medication that also prevents migraines, as opposed to being on a separate anti-hypertensive and medication for migraine prophylaxis. Other predictable queries include interpretation of lab results and normal lab values (Cimino, 2006). In general, finding this sort of information is not challenging, but navigating to it is tedious.

Our next step was to create a clinical vignette to help us develop use cases. We chose the primary care context, both because it is the most common type of physician visit and because Professor Del Fiol and I have both practiced primary care. Arguably, primary care sees the broadest range of illnesses and may have the largest information needs. I chose the disease of rheumatoid arthritis (RA) because it has a number of new treatment options[2] ("biologic agents"), all of which have been approved since I completed

[2] In fact, most of the targets for these medications had not been discovered when I was in medical school. We did not have a dedicated Immunology course, so it is not just that I am unfamiliar with the medications, but also that I was never exposed to the basic science that underlies these medications.

residency. I have had no formal education regarding any drugs in this class, and I do not prescribe them in the Emergency Department.

Our vignette was:

> *John Rheumatology is a 59 y/o WM with a history of **rheumatoid arthritis** (RA), **hypertension** (HTN), **obesity**, and chronic **right knee pain** (presumably due to a combination of RA and osteoarthritis of the knee, aggravated by obesity). His rheumatoid arthritis has been managed with **methotrexate** (MTX), which is a Disease Modifying Anti-Rheumatic Drug (DMARD). He presents to the office today because his knee pain is getting worse and he is also having some hip and wrist pain. He has been receiving advertisements in the mail from manufacturers of a variety of other medications for rheumatoid arthritis and he wonders if he might benefit. However, reading through the potential side effects (cancer, sudden death), he is appropriately hesitant to try these medications if there is no clear benefit. While you do not anticipate prescribing a biological agent without referral to a rheumatologist, you would still like to be able to give John some general information about the risk:benefit ratio of using these medications and you wish to supply him with some questions to ask the rheumatologist, if he chooses to be evaluated.*

Based on this vignette, I developed a series of clinical questions that I thought could plausibly arise during the course of this visit

- What Disease Modifying Anti-Rheumatic Drugs (DMARDs) are available other than methotrexate (MTX) and biologic agents? Is there any utility in older DMARDs, e.g., gold salts?
- There are a number of new biologic agents that have been approved for the treatment of rheumatoid arthritis including TNF-α inhibitors, interleukin-1 inhibitors, interleukin-6 inhibitors, CD-28 blocking agents, and B-cell depleting agents which target CD-20. Are certain classes superior to others? Are certain drugs within each class superior? Are they typically prescribed as monotherapy or in combination with another drug?
- What are the criteria for determining "treatment failure" or "inadequate response" to current therapy in this patient?
- Is a combination of pharmacologic and non-pharmacologic treatments superior to either alone (e.g., would knee joint replacement provide more relief than medication)?
- Are there any special considerations when looking at both the age and the gender of the patient?
- Does the patient's underlying hypertension change the treatment of his RA?
- Are there any studies or articles looking at a patient population similar to our patient (RA + HTN, RA activity not controlled on MTX)? In other words, what is the closest comparable cohort that has been studied?

As noted in the pilot study, we wanted to provide compact answers to questions like those listed above. It is possible that these answers are available and can be used *verbatim*

from existing resources. If so, we would like to provide them to the clinician. If not, we would like to reformat the information that is present in a way that does not change the initial meaning and still suggests a clear course of action. Tables and images that are highly informative are also acceptable. I was asked by Professors Mostafa and Del Fiol to mock up a presentation layer based on the premise that the underlying information retrieval would be perfect. As noted in the CDS use case diagram (Figure 4), electronic CDS always competes with either a formal consult (a referral) or an informal ("curbside") consult with a colleague. If consultants are readily available, physicians are less likely to spend time searching for information (which is not to say that they will get their question answered by the consultant.) Therefore, the CDS system must provide useful information very rapidly.

Given those general specifications, I proceeded to complete an informal requirements analysis of the system's user interface. Requirements analysis is a process of discovering the boundaries of a software project (the scope), what a software system is supposed to do (the software requirements specification), and how users are going to interact with the system (the user requirements). While there is considerable variation in how this is accomplished, the general framework is well established.

The scale of the requirements analysis process is dependent on the scale and complexity of both the business and the software. In the case of clinical decision support systems and healthcare, both the business and the software are highly complex. Medical software systems also have to contend with rapidly changing standards of care and contingencies

which are, by definition, unpredictable (new diseases, natural disasters, unpredictable financial incentives, and arbitrary regulations). This means that the requirements for a medical software system cannot be completely specified prior to the start of coding, and it can be expected that software requirements are likely to change frequently, even during the course of trying to characterize the requirements of the system. It also means that medical software, particularly CDS systems, are never finished, but rather in a continual state of development. Therefore, traditional waterfall techniques for software planning are a poor fit for medical software systems. Various software engineering techniques that rely on small, "completable" steps and rapid prototyping are preferred when dealing with medical software. Finally, medical software needs to be highly reliable, as it may harm patients if it fails.

The usual steps in the requirement analysis process are as follows:

Determine the scope and general vision for the project, usually via a discussion with the business management:

- What is the business case for this system?
- What are the boundaries of this project?
- Who are the stakeholders?

In the general case of CDS systems, the vision is to produce cheaper and better care for patients by improving the selection of therapies, which should produce better outcomes, fewer "medical misadventures," and lower costs. The boundaries of this project are that this particular CDS system is limited to the selection of therapies, and it is confined to

retrieving and displaying situationally relevant information It is assumed that the patient already has a diagnosis and that the diagnosis is correct with a realistic degree of certainty. My portion of the project is specifically confined to the interaction and interface that the clinician user will confront. The stakeholders for this project are the clinician users of the interface, and I am specifically targeting primary care physicians. In a production system, there are many indirect stakeholders, starting with patients, but also including the public, payers, other providers, and producers (the 5 Ps of health data systems: http://www.healthdatainnovation.com/category/tags/5-ps).

Determine the user requirements for the system, usually via user/stakeholder interviews and observation.

These interviews are used to discover what tasks the users are trying to accomplish (as opposed to what the users think the system should do). The information gathered during these interviews is used to produce use cases (traditional requirements analysis) or user stories (agile). In the general case of all CDS systems, the direct stakeholders are numerous and include (at a minimum) patients, all ancillary and nursing staff, payors, facility administrators, government agencies, malpractice insurers, pharmacists, pharmaceutical companies, medical device manufacturers, and providers. For the purpose of the "Knowledge Summary" CDS system, however, the users are the providers. The use case is the entry and review of patient data in the EHR either during the visit or shortly before the visit. The information in the knowledge summary should be tailored to the situation (a clinic visit), the specialty of the provider (primary care), the experience of the provider (attending physician), and the number of patients with this

specific disease on the provider's panel (less than 5/2300 in this practice). In this instance, I am a member of both the design group and the user group. This allows for extremely rapid design iteration and deep understanding of how this software would be used, although it is certainly not a consensus view.

Determine the functional requirements of the system (i.e., what the system needs to do, not how it is going to do it.)

Ideally, this system should provide the same guidance that a trusted specialist would provide if he/she were in the room, looking over the primary care physician's shoulder. Above and beyond providing actionable answers to questions on demand, a consultant could also suggest courses of action not considered by the primary care physician or even anticipate suboptimal decisions that the physician might make (e.g., "Well, the guidelines suggest that you could add a TNF inhibitor at this point, but I can tell you that guys like this get a lot more durable relief from a knee replacement.") The goal is for the system to provide concise, germane, and actionable clinical information to the clinician at the time that the information is needed (Fischer, 2012). While this system will be implemented as "passive" CDS, we hope to overcome the major physician usability issues present in many "clinical reference"-style CDS systems, namely that it takes too long to search and navigate them. Additionally, although not a specific complaint of physicians, we would also like to address the issue of "unknown unknowns" by pushing information to clinicians that they were unaware that they needed. Although doctors are subjected to a great deal of unsolicited material in the form of journal articles, it is rarely presented at the time it is needed.

An additional necessary functional characteristic of the system is that it should supply the best quality information available, and the quality of the information presented should be readily evident. While the term "quality" has the potential to be subjective, in this context it means the quality of evidence, using criteria established by experts in evidence-based medical practice (Howick et al., 2009).

Recognizing that most questions in medicine do not have high-quality evidence available to give definitive answers, other factors that physicians often rely upon, such as the authority and reputation of both the authors and the journal, the potential clinical impact (small effects for non-serious illness should be weighed less than large effects), the currency of the information, trial size (if it is a clinical trial), and potential sources of bias, should be weighed as part of the "quality" measure. The components of the quality metric should be readily apparent (Shurtz & Foster, 2011). If there is no information available, the interface should indicate that, and let the physician reach that conclusion both quickly and decisively.

Additionally, local information should be incorporated into the knowledge summary (for example, "Is a stress echo, myocardial PET scan, or a nuclear medicine stress test the preferred cardiac imaging test *at this specific hospital?*")

Last of all, the interface needs to be *fast* and highly interactive. In practical terms, this means that the client will need to pre-fetch a significant amount of information and store

it locally to achieve rich interactivity when reviewing content. All information should be fetched prior to interaction with the user.

Non-functional requirements - as noted above, the prevalence of commercial EHRs has made it imperative that the system be interoperable via some common standard (e.g., openinfobutton).

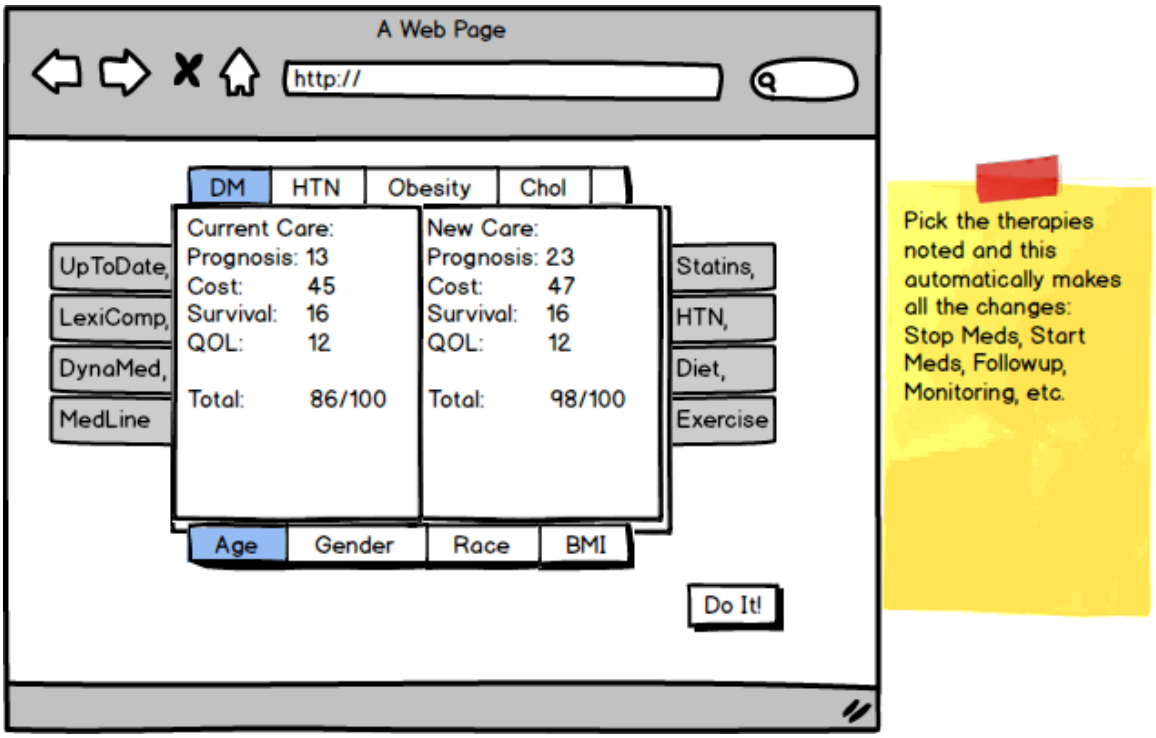
It should function on Epic, Cerner, and other major vendors, and the user interface should adhere to the user interface conventions already present on the system. For example, <F8> closes a window in Epic, *ergo*, <F8> should also close the CDS window.

User Interface – Iteration 1

Adjusting the query by selecting patient characteristics or treatments

My first mockup of the UI focused on using the "patient characteristics as the query." Data elements that are both universally present and typically well maintained (and therefore used for CDS) in most EHRs include the medication list, allergies, problem list, past medical history, age, and gender. Additionally, I wanted to provide some sort of "quality metric" that was easily digestible by the physician so that different courses of action could be compared. In Figure 5, the tabs on the top represent patient comorbidities and the tabs across the bottom represent demographic variables. If I wanted to compare my current patient to a group that contained only men, I would de-select the gender (= male). By toggling the tabs along the upper and lower portions of the center panels, I could select a different cohort of patients as the comparison group. Selection of different

information resources along the left edge would use the information in those resources to calculate the outcomes. Tabs along the right side represent different interventions that could be tried on this patient. The choices made here automatically feed back to the



orders section of the record.

Figure 5 - Output of a quantitative score based on changing therapies selected with the right tabs.

Figure 6 shows another interface for selecting medications. In this version, I have added a "Fast" and a "Flex" page. The "Fast" page will have a fairly static display of highly informative sentences and medication doses. The "Flex" page will allow more exploration and have links out to full search engines and knowledge resources. However, it will still be optimized for displaying actionable, factual, clinical items. If more than 1

to 2 minutes of review are needed, the user is likely to be better served by a full search interface to PubMed or UpToDate.

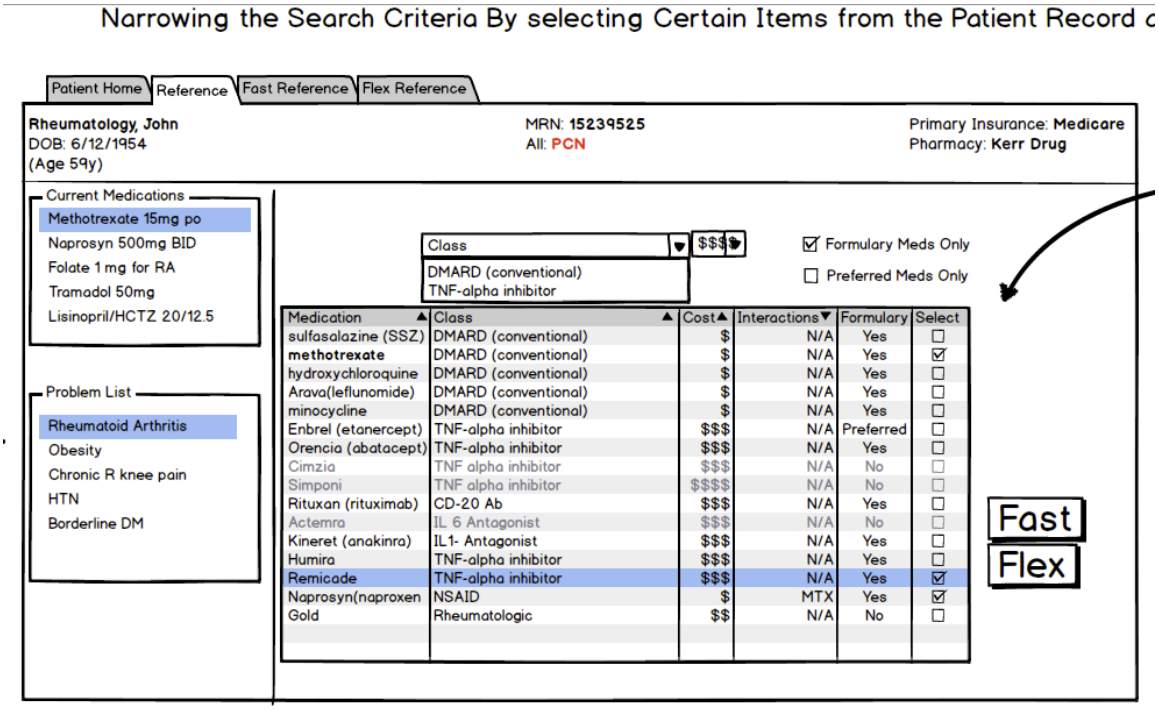


Figure 6 - Another way to use the problem list, past medical history, and medications to construct a query. The panels in the center give more details about the medications, including cost and formulary. Formulary is one form of local information that influences therapy choices.

User Interface – Iteration 2

3 columns with current treatment in the center

The expected interaction in the first iteration (Figure 5) with regard to changing therapies was not very clear, and I felt that Figure 6 displayed too much detail (e.g., cost) prior to

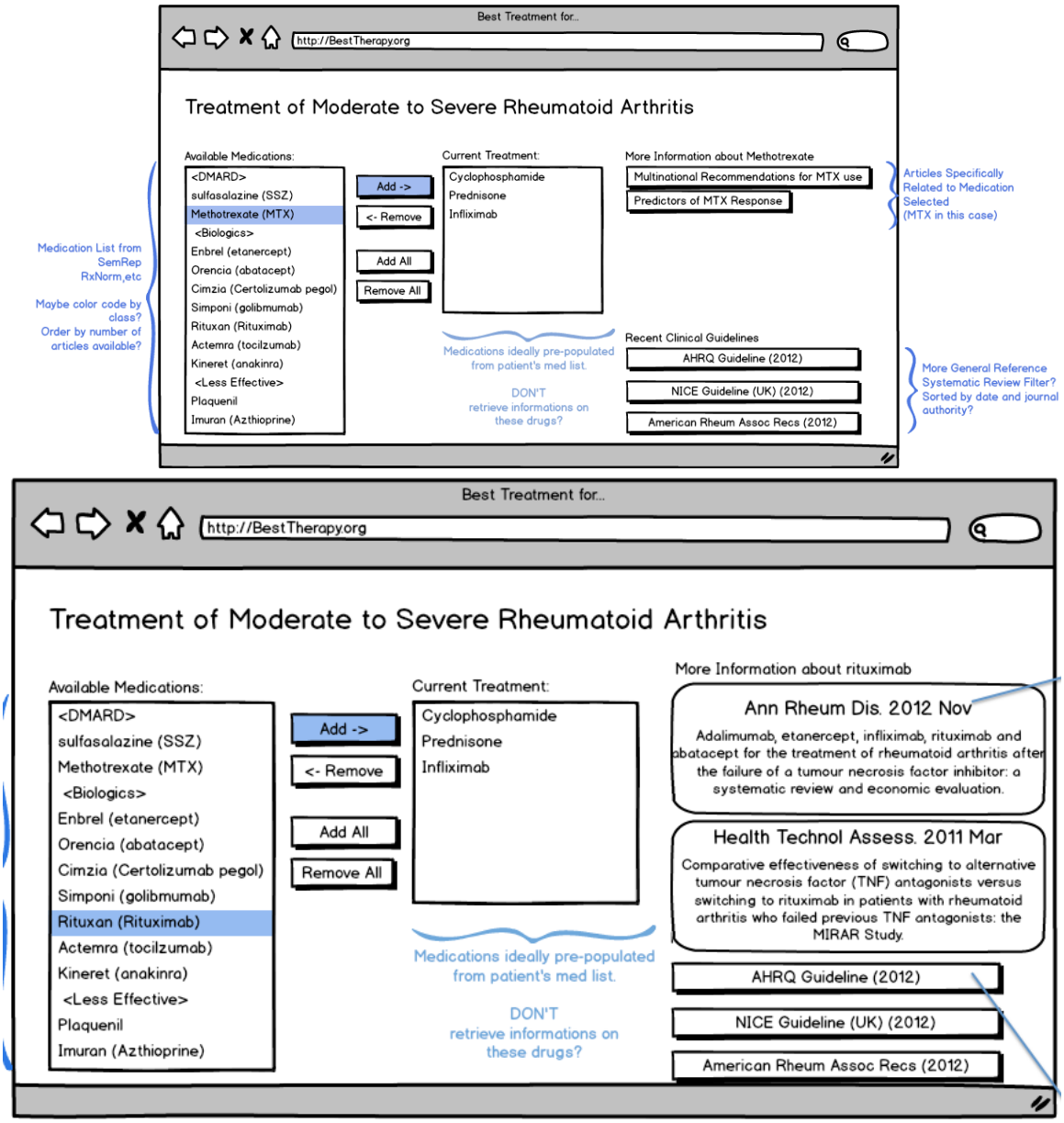


Figure 7 – Top and bottom panels illustrate the interaction when different medications are selected (top: MTX; bottom: Rituximab). Selection of medications in the left panel drives the display of medication-specific clinical trial information.

really evaluating all of the clinical factors. While cost is definitely an issue, other issues should take precedence (e.g., safety and effectiveness.)

This iteration (Figure 7) has a more "normal" flow from left to right. It also has the current medications appearing in the center of the screen with a very conventional set of "Add/Remove" buttons that clearly indicate what to expect when you press them. The right-hand column will change dynamically, depending on the medications chosen. I had to remove the demographics and co-morbid diseases from this interface because there was not enough room, particularly with this prototyping tool (Balsamiq), which is not ideal for producing high-fidelity prototypes.

As noted above in the pilot study, document level retrieval was too coarse. We chose to use a sentence-level index of all PubMed abstracts (Kilicoğlu et al., 2008); by using the PubMed clinical queries filter, followed by the Semantic MEDLINE index, we evaluated sentences that were related to the treatment of rheumatoid arthritis. I also went through the UpToDate articles on the treatment of moderate to severe rheumatoid arthritis and picked out sentences that seemed to summarize the most important facts about the treatment of rheumatoid arthritis. I then considered how these would be displayed.

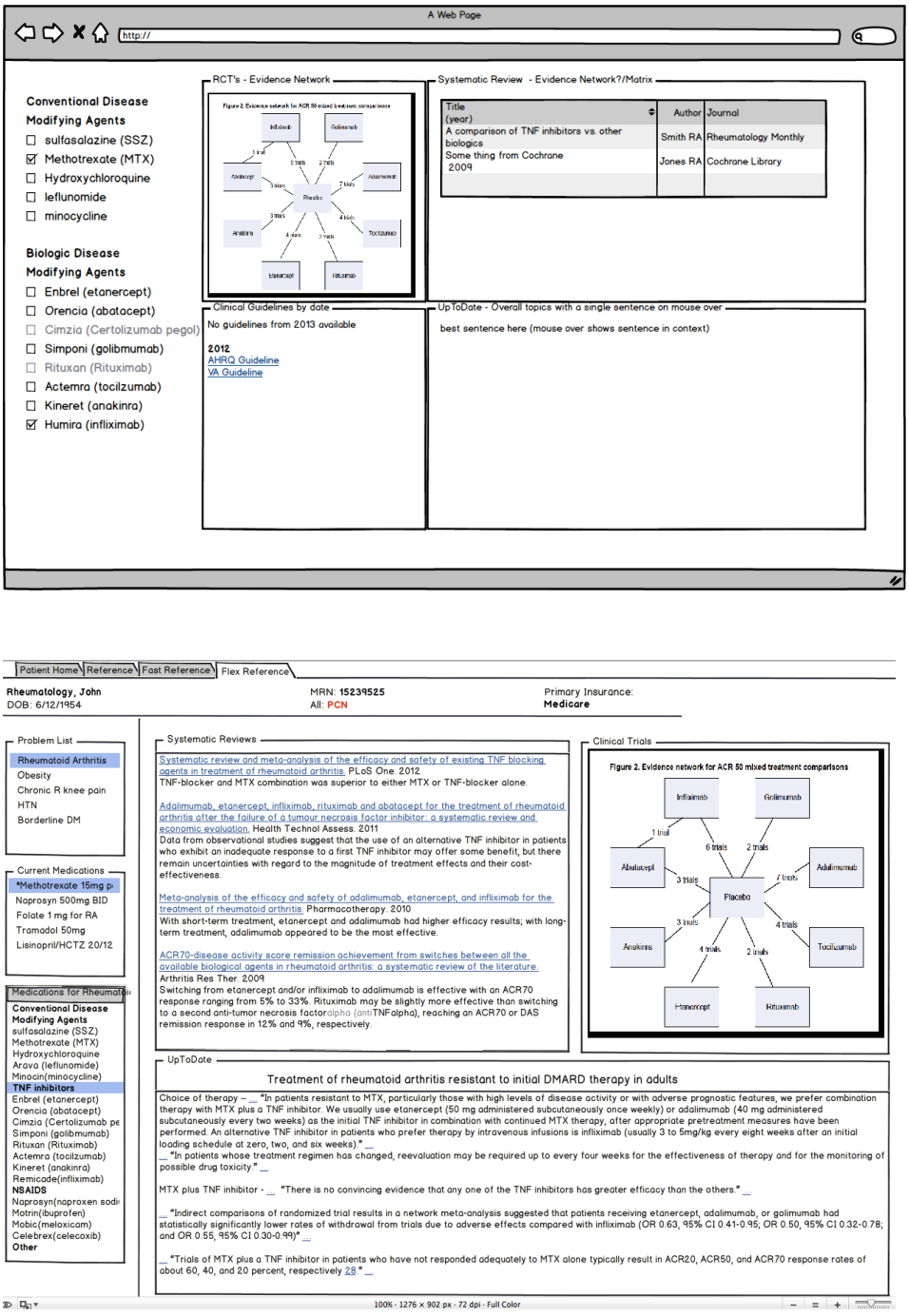


Figure 8 - Top Panel - I needed more space for content and less for selection of medication. Bottom Panel - I had to make a decision about cluttering up the screen and slowing comprehension and scanning of the first page, or requiring additional clicks for navigation. I decided that I preferred fewer clicks.

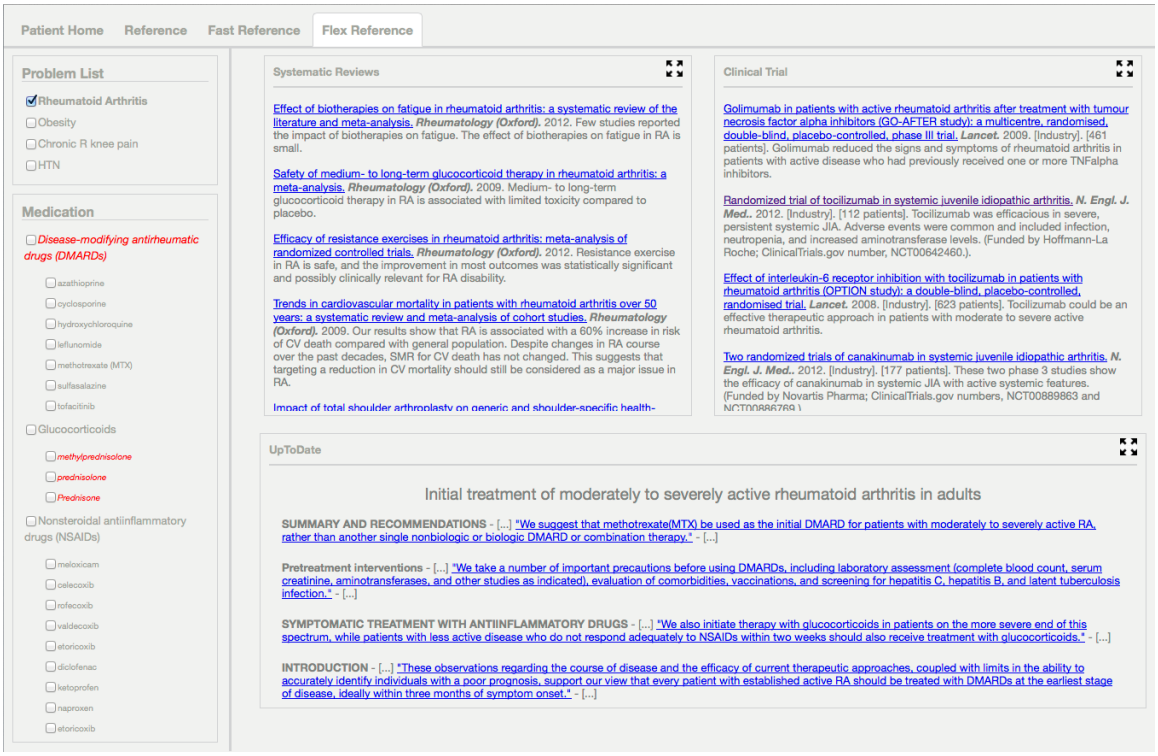


Figure 9 - Current high-fidelity prototype (http://miksa.ils.unc.edu/ksvis/). High-yield sentences are available on the home page. Mousing over content in the right panel reveals information such as pharma sponsorship, trial size, and journal publication, all without having to navigate away from the page. Selection of medication and co-morbidities in the left column changes the information retrieval in the right-hand panels. Each panel retrieves a different kind of clinical content (Systematic Review, UpToDate, or Clinical Trials.)

Conclusion

Most clinical questions can be answered using existing clinical reference material, but it is frequently difficult for physicians to find that information in a timely fashion. The resources that contain the answers are heterogeneous and include sources such as institutional FAQs, government sponsored systematic reviews, citation databases and commercially licensed electronic textbooks. Additionally, access is also likely to vary between healthcare organizations, particularly for commercially licensed sources. It is unlikely that these resources will ever be unified into a single repository as they are competitors in the marketplace; therefore, a software agent that has the ability to conduct a federated search and then summarize clinical reference materials will remain a required capability for enabling a clinician's daily work.

As noted in the introduction, Ely's extensive work into the specific kinds of questions that physicians ask, combined with his thorough documentation of the actual barriers to information retrieval and use, was critical to informing the design of this user interface. In particular, his observation that physicians consciously decide not to pursue clinical questions suggests that current user interfaces require too much time investment for the amount of knowledge gained. Therefore, minimizing physician search effort by prediction of physician information needs would be a major advance. This "it needs to be worth the effort" behavior was originally described as "information foraging" theory –

namely, that people will return to resources that have a high density of useful information (Pirolli, 1999).

When Ely first collected the questions, he was unable to find answers to many of the questions that physicians asked. Ely's decision to make the questions he collected available was invaluable in allowing me to decide if adequate content had been developed for physician clinical reference over the course of the past decade. My ability to answer Ely's original questions using current content suggests that both content and indexing have become more complete over the last decade.

Professors Mostafa and Del Fiol invited me to assist them with the development of an improved user interface for clinical reference material as a subject matter expert. Del Fiol had extensive "real-world" implementation experience with Infobuttons in production EHRs at both Intermountain Health and the Veterans Administration. Additionally, he was one of the major authors of the HL7 Infobutton interface. However, he felt that there was no underlying theory guiding the direction of the design of the next generation of clinical reference interface. The feedback he had gotten from users was that the medication and lab references worked well, but that other parts of the Infobutton system did not provide adequately targeted information. In particular, queries sent to both PubMed and commercial vendors tended to return either large numbers of topically relevant, but clinically irrelevant results or a single long document that was both topically and clinically relevant, but whose length made it difficult to find the specific clinically relevant content. His goal was to return results that were both comprehensive and terse.

Mostafa had extensive experience in developing interfaces for both provider and consumer health information retrieval, including retrieval of non-text documents such as MRI scans. More importantly, however, he had an underlying theoretical understanding of how people approach searching and satisfying information needs. Based on Mostafa's guidance, a number of potentially applicable information-seeking theories were reviewed, and the "information foraging" framework was selected to guide the development of the user interface.

My participation in the initial pilot study was focused on using current resources to answer clinical questions that were originally asked by physicians during the course of a typical day in an outpatient primary care office over a decade ago. Overall, I felt that the questions in the Clinical Questions collection were still representative of the questions that physicians ask, which upon reflection is unsurprising, as the overall types of illnesses have not changed in the last decade, even if the treatments have.

I used reference sources that were not part of my "normal" search technique (e.g., Lexicomp instead of Epocrates, DynaMed instead of UpToDate, Google Scholar instead of PubMed), which exposed me to the task of learning some new user interfaces. Additionally, my specialty is Emergency Medicine and while there is some overlap with primary care, the information that I knew at baseline for many of these questions was not enough to yield an appropriate answer for the clinical situation of working in a primary care office. Nonetheless, the categories of information that I use to determine if a retrieved item is clinically relevant, authoritative and unbiased are likely to be similar

across specialties. For example, when reviewing actual clinical trials, I need to know details about the patient population of the trial (how many patients, what country, how long ago, characteristics that are markedly different from the patient at hand) and about the researchers (conflicts of interest in terms of specialty, pharmaceutical ties, and previous publications) to inform my expert opinion. The current iteration of the user interface addresses these issues admirably.

When considering Ely's observations about when physicians choose not to pursue answers to clinical questions, I considered my own practice of Emergency Medicine. I am frequently confronted with making the choice of what information to pursue, both in the patient medical record and in clinical reference sources. Many of the patients I see are both diagnostic and treatment puzzles who have undergone extensive diagnostic workups and treatment protocols prior to abandoning their local healthcare providers and presenting to the UNC Emergency Department. The probability that these patients have an easily diagnosable problem with an appropriate, evidence-based treatment is very low. The uncertainties of actual medical practice very much informed my interface design. The ideal interface would give me the same confidence in my decision as a "curbside" consult with a trusted colleague, including confirmation that there is no research data about the current patient's condition, and that treatment decisions must be made based on expert opinion.

This project changed my perception of resident and physician search behavior. As I was working on this project, I became very sensitive to how residents searched for

information. While I was initially surprised that physician trainees would accept "non-professional" sites as authoritative, I was astounded when I discovered that they continued using the same search strategies when it was suggested that they needed to use an authoritative resource, presumably because the UI of the authoritative resource was either unfamiliar or provided a poor user experience.

The general choice of "convenient" over "authoritative" was illustrated when I walked into a patient's room and found an intern using an online translation dictionary to communicate with a patient who only spoke a foreign language, despite the immediate availability of certified medical translators via a telephone service. Nonetheless, use of the online translation system saved a small amount of time, and therefore it was selected by the intern because he perceived that there was not enough time to use the translation service. Therefore, it is critical that the Knowledge Summary user interface not only contains the correct answer, but it must also contain affordances that encourage use. In other words, it requires a design that immediately rewards the user for the interaction and does not require learning new skills to navigate at a basic level.

In conclusion, the production of clinical reference material for physicians is lucrative, competitive, and time-consuming. In the era of printed textbooks, the major tool for content production was a pen and paper, or possibly a word-processor, both tools with which physicians are familiar and this enabled physicians to be content producers. With new media, however, the ability to produce content, particularly interactive content, has become a specialized skill and most physicians no longer have the ability to produce this

sort of material. While web page production has been viewed by traditional medical textbook publishers as the work that occurs late in the content production process, it is clear that it needs to occur much earlier if the final product is expected to produce a good user experience. Developers must become the content developers, but they do not have the requisite intuition regarding content that a physician would want to see or how they would want it presented. Unlike consumer-oriented content, where the developers have at least some passing familiarity with the role of being a content consumer, the content produced for physicians requires at least a decade of formal education and clinical experience to appreciate. As the medical textbook publishers have discovered, simply placing static HTML copies of textbooks online does not work.

Textbooks that were "born online" have quickly surpassed established printed materials, both in terms of popularity with physicians and (presumably) revenue for publishers e.g., Harrison's Principles of Internal Medicine is published once every three years and costs $125 for the e-book. A subscription to UptoDate costs $499 per year. The key advantage that UpToDate has over traditional printed textbooks is that it possesses significantly improved navigation, via hyperlinks, between related content. This navigation is not replicated by traditional indexing. While there are other online widely available clinical reference sources (e.g. WebMD, eMedicine.com, MDConsult), I believe that there are three design decisions that result in UpToDate's popularity. The first is that it was explicitly designed as an online resource (rather than trying to repackage printed material). The second is that it has maintained "pure" professional-level content, rather than trying to cater to other audiences. The last is that the revenue model is entirely based

on subscriptions and the user interface does not contain extraneous advertising material, which I feel has a significant adverse effect on my user experience. As predicted by information foraging theory, physicians are willing to pay for content that efficiently and consistently satisfies information needs.

Knowledge Summaries are an example of the next wave of clinical reference content. They require much more developer input than resources such as UpToDate (which is essentially a collection of static webpages, despite extensive crosslinking and frequently updated content.) Knowledge Summaries are dynamically generated from constantly evolving content, and their final form will depend on the rich-context information that is pulled from electronic health records, context information that is both about the patients and the physicians treating them. This will present significant challenges for developers because Knowledge Summaries will require deep understanding of physician workflows and thought processes, as well as knowledge about medical science itself. It is impossible for developers to acquire this knowledge on their own without practicing medicine, so they will require constant feedback from practicing physicians. Similarly, physicians who are involved in the development process must possess special knowledge regarding cognitive psychology, information science, and user interface design in order to effectively communicate systems requirements to developers. Software systems that are as complex as the Knowledge Summaries system will not be produced by anything less that expert software engineers and developers. In other words, it will require a team that is composed of both developers and clinicians working closely together to create a reliable and useful product. Last of all, because medicine is constantly changing, there

will need to be a continuous content development process that can respond to changes in medical practice.

In order to facilitate the continuous content development process, Web 2.0 techniques for incorporating both actively and passively collected user feedback. This will result in a content system with a constant, real-time quality control process that is unlike anything in medical publishing today. How to keep users engaged in the feedback loop will be a challenge, however. While it is unknown what degree of participation will be required to maintain accurate content, it can be assumed that there will need to be something more compelling that community spirit to encourage content validation and production. One mechanism that is utilized by UpToDate is rewarding physicians with Continuing Medical Education credit for supplying written feedback about searches and the content they reviewed. This is a very effective mechanism of promoting physician participation.

The Knowledge Summary interface is still under development (NIH Grant 1R01LM011416-01. There are a number of innovative features in the interface: the incorporation and annotation of heterogenous sources makes it very easy to find information about the less-tangible measures of research quality, such as pharmaceutical company sponsorship and clinical trial size. It is also the first alternative interface for commercial knowledge sources, such as UpToDate. It also accommodates the "real world" lack of evidence for most clinical decisions by having dynamic criteria for what constitutes quality clinical reference material. Overall, it is much more representative of

the type and quantity of information that physicians would “retrieve” by consulting a colleague.

There are innumerable clinical variables that are considered by clinicians when evaluating patients and their own information needs (Table 2). It is not hard to imagine how specific information needs might be predicted for each patient-provider dyad, given enough data and programming effort. For example, a physician seeing a patient with a disease that the physician rarely encounters might need more information than a physician seeing a patient with a disease that they treat on a regular basis. Some diseases are so rare that it is likely that very few physicians are experts on the disease, so it might be worth providing that information to all providers who encounter a particular patient for the first. Additional key capabilities for a future iterations of a Knowledge Summary system would be the recognition of "new" information (or at least "new to this specific clinician”, as the question “Are there any changes in management of disease X?” is a common motivation for seeking information (as opposed to a specific question type) . Future research will focus on which of these variables is useful for predicting physician information needs.

Table 2 - Characteristics of patients, providers, and facilities that might be predictive of information needs for a particular encounter.

Patient	Facility	Provider	Interactions
Age Gender Comorbidities Previous Medical Management Errors Recent Lab Results Recent Imaging Results Other Test Results Recent Consults Secondary Gain (Worker's Comp, personal injury liability) Are this patient's complaints predictive of illness?	Type (Hospital vs. Clinic vs. Nursing Home) Load (Routine vs. Busy vs. Disaster) Specialty Care Facility (Pediatric, Cardiac, Oncology, Burn Hospital) Resources (Lab, Imaging, Consultants, Ancillary Services, Financial)	Years in Practice (Intern vs. Attending) Years since formal education Exposure to Peers (e.g., rural solo practice vs. academic medical center) Typical Patients Similar Cases Previous Medical Management Errors Previous Malpractice Claims Fatigue	Does this physician usually see this type of patient in this facility? Does the physician know this patient? Is this at 4:00AM on a Monday in the ED or at 10:00AM on a Tuesday at the office? Other local effects – is a certain disease endemic to this geographic area or to this patient population. Is this a VIP? Has this patient been seen by other providers for the same problem?

Bibliography

Abbas, J., Schwartz, D. G., & Krause, R. (2010). Emergency medical residents' use of Google® for answering clinical questions in the emergency room. *Proceedings of the American Society for Information Science and Technology, 47*(1), 1–4. doi:10.1145/1882992.1883049

Abbott, K. H. (2001). Chapter 9: human factors engineering and flight deck design. In *The avionics handbook*. Boca Raton, FL: CRC Press.

Bergus, G. R., Emerson, M., Reed, D. A., & Attaluri, A. (2006). E-mail teleconsultations: well formulated clinical referrals reduce the need for clinic consultation. *J Telemed Telecare, 12*(1), 33–38. doi:10.1258/135763306775321434

Bright, T. J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., Coeytaux, R. R., Samsa, G., Hasselblad, V., Williams, J. W., Musty, M. D., Wing, L., Kendrick, A. S., Sanders, G. D., & Lobach, D. (2012). Effect of clinical decision-support systems: a systematic review. *Annals of Internal Medicine, 157*(1), 29–43. doi:10.7326/0003-4819-157-1-201207030-00450

Chen, E. S., & Cimino, J. J. (2003). Automated discovery of patient-specific clinician information needs using clinical information system log files. *AMIA Annu Symposium Proc*, 145–149.

Cimino, C., & Barnett, G. O. (1991). Analysis of physician questions in an ambulatory care setting. *Proc Annu Symp Comput Appl Med Care*, 995–999.

Cimino, J. J. (2006). Use, usability, usefulness, and impact of an Infobutton manager. *AMIA Annu Symp Proc*, 151–155.

Cimino, J. J., Johnson, S. B., Aguirre, A., Roderer, N., & Clayton, P. D. (1992). The MEDLINE button. *Proc Annu Symp Comput Appl Med Care*, 81–85.

Cobin, R. H. (2002). Subspecialist care improves diabetes outcomes. *Diabetes Care, 25*(9), 1654–1656.

Collins, S., Bakken, S., Cimino, J. J., & Currie, L. (2007). A methodology for meeting context-specific information needs related to nursing orders. *AMIA Annu Symp Proc*, 155–159. PMCID: PMC2655787

Cook, D. A., Sorensen, K. J., Wilkinson, J. M., & Berger, R. A. (2013). Barriers and decisions when answering clinical questions at the point of care: a grounded theory study. *JAMA Internal Medicine*. doi:10.1001/jamainternmed.2013.10103

Covell, D. G., Uman, G. C., & Manning, P. R. (1985). Information needs in office practice: are they being met? *Annals of Internal Medicine, 103*(4), 596–599.

Dhaliwal, G. (2013). Known unknowns and unknown unknowns at the point of care. *JAMA Internal Medicine*. doi:10.1001/jamainternmed.2013.7494

Del Fiol, G., Huser, V., Strasberg, H. R., Maviglia, S. M., Curtis, C., & Cimino, J. J. (2012). Implementations of the HL7 context-aware knowledge retrieval ("Infobutton") standard: challenges, strengths, limitations, and uptake. *J Biomed Inform, 45*(4), 726–735. doi:10.1016/j.jbi.2011.12.006

Eaton, C. B., Murphy, J. B., & Hunt, V. R. (1997). Outcome of acute myocardial infarction according to the specialty of the admitting physician. *The New England Journal of Medicine*, *336*(22), 1608–author reply 1608–1609.

Ely, J. W. (2004). Answering physicians' clinical questions: obstacles and potential solutions. *Journal of the American Medical Informatics Association, 12*(2), 217–224. doi:10.1197/jamia.M1608

Ely, J. W., Osheroff, J. A., Ebell, M. H., Chambliss, M. L., Vinson, D. C., Stevermer, J. J., & Pifer, E. A. (2002). Obstacles to answering doctors' questions about patient care with evidence: qualitative study. *BMJ: British Medical Journal*, *324*(7339), 710.

Ely, J. W., Osheroff, J. A., Gorman, P. N., Ebell, M. H., Chambliss, M. L., Pifer, E. A., & Stavri, P. Z. (2000). A taxonomy of generic clinical questions: classification study. *BMJ*, *321*(7258), 429–43.

Ely, J. W., Osheroff, J. A., Maviglia, S. M., & Rosenbaum, M. E. (2007). Patient-Care Questions that Physicians Are Unable to Answer. *Journal of the American Medical Informatics Association, 14*(4), 407–414. doi:10.1197/jamia.M2398

Ely, J. W., Osheroff, J. A., Ebell, M. H., Bergus, G. R., Levy, B. T., Chambliss, M. L., & Evans, E. R. (2000). Analysis of questions asked by family physicians regarding patient care. *The Western Journal of Medicine, 172*(5), 315–319.

Fischer, G. (2012). Context-aware systems: the "right" information, at the "right" time, in the "right" place, in the "right" way, to the "right" person. Presented at the AVI '12: *Proceedings of the International Working Conference on Advanced Visual Interfaces*, ACM Request Permissions. doi:10.1145/2254556.2254611

Goldstein, L. B., Matchar, D. B., Hoff-Lindquist, J., Samsa, G. P., & Horner, R. D. (2003). VA stroke study: neurologist care is associated with increased testing but improved outcomes. *Neurology, 61*(6), 792–796.

Graber, M. A., Randles, B. D., Ely, J. W., & Monnahan, J. (2008). Answering clinical questions in the ED. *The American Journal of Emergency Medicine, 26*(2), 144–147. doi:10.1016/j.ajem.2007.03.031

Howick, J., Chalmers, I., Glasziou, P., Greenhalgh, T., Heneghan, C., Liberati, A., et al. (2011). *The Oxford 2011 levels of evidence*. Retrieved from http://www.cebm.net/index.aspx?o=5653 on 11/4/2013.

Jerome, R. N., Giuse N. B., Gish K. W., Sathe N. A., & Dietrich M. S. (2001). Information needs of clinical teams: analysis of questions received by the clinical informatics consult service. *Bulletin of the Medical Library Association, 89*(2), 177–184.

Jonnalagadda, S. R., Del Fiol, G., Medlin, R., Weir, C., Fiszman, M., Mostafa, J., & Liu, H. (2013). Automatically extracting sentences from MEDLINE citations to support clinicians' information needs. *Journal of the American Medical Informatics, 20*(5), 995–1000.

Kilicoğlu, H., Fiszman, M., Rodriguez, A., Shin, D., Ripple, A. M., & Rindflesch, T. C. (2008). Semantic MEDLINE: a web application for managing the results of PubMed searches. *Proc Third Int'l Symposium for Semantic Mining in Biomedicine*, 69–76.

Krause, R., Moscati, R., & Halpern, S. (2011). Can emergency medicine residents reliably use the Internet to answer clinical questions? *Western Journal of Emergency Medicine 12*(4), 1–8. Retrieved from http://escholarship.org/uc/item/1fn2761x;jsessionid=84FB0552CF713B18EC58757971578DEA

Krumpholz, A. (2012). Automated medical literature retrieval. *Australasian Medical Journal 5*(9), 489–496. doi:10.4066/AMJ.2012.1378

McClellan, M. B., McGinnis, J. M., Nabel, E. G., Olsen, L. M., Institute of Medicine. (2008). *Evidence-based medicine and the changing nature of health care*. Washington, DC: National Academies Press.

Mendonça, E. A., Cimino, J. J., & Johnson, S. B. (2002). Using patient data to rank records of literature retrieval. *Proc AMIA Symp*, 1103.

Meyer, A. N. D., Payne, V. L., Meeks, D. W., Rao, R., & Singh, H. (2013). Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study. *JAMA Internal Medicine*. doi:10.1001/jamainternmed.2013.10081

Nash, K. B., Josephson, S. A., Sun, K., & Ferriero, D. M. (2013). Should there be pediatric neurohospitalists? *Neurology, 80*(10), 957–962. doi:10.1212/WNL.0b013e3182840bd3

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review, 106*(4), 643.

Pluye, P. Grad, R. M., Dunikowski, L. G., & Stephenson, R. (2005). Impact of clinical information-retrieval technology on physicians: a literature review of quantitative, qualitative, and mixed methods studies. *International Journal of Medical Informatics, 74*(9), 745–768.

Shurtz, S., & Foster, M. J. (2011). Developing and using a rubric for evaluating evidence-based medicine point-of-care tools. *Journal of the Medical Library Association: JMLA 99*(3), 247–254.

Simon, H. A. (1969). Designing organizations for an information-rich world. In Martin Greenberger (Ed.), *Computers, communication, and the public interest* (pp. 40–41). Baltimore, MD: The Johns Hopkins Press.

Smith, A. (2003). Expertise in practice: an ethnographic study exploring acquisition and use of knowledge in anaesthesia. *British Journal of Anaesthesia 91*(3), 319–328.

Weiner, J. P., Kfuri, T., Chan, K., & Fowles, J. B. (2007). "e-Iatrogenesis": the most critical unintended consequence of CPOE and other HIT. *Journal of the American Medical Informatics Association: JAMIA, 14*(3), 387–8; discussion 389. doi:10.1197/jamia.M2338