Research Paper ■

Using Wireless Handheld Computers to Seek Information at the Point of Care: An Evaluation by Clinicians

Susan E. Hauser, PhD, Dina Demner-Fushman, MD, PhD, Joshua L. Jacobs, MD, Susanne M. Humphrey, MLS, Glenn Ford, George R. Thoma, PhD

Abstract **Objective:** To evaluate: (1) the effectiveness of wireless handheld computers for online information retrieval in clinical settings; (2) the role of MEDLINE® in answering clinical questions raised at the point of care.

Design: A prospective single-cohort study: accompanying medical teams on teaching rounds, five internal medicine residents used and evaluated MD on Tap, an application for handheld computers, to seek answers in real time to clinical questions arising at the point of care.

Measurements: All transactions were stored by an intermediate server. Evaluators recorded clinical scenarios and questions, identified MEDLINE citations that answered the questions, and submitted daily and summative reports of their experience. A senior medical librarian corroborated the relevance of the selected citation to each scenario and question.

Results: Evaluators answered 68% of 363 background and foreground clinical questions during rounding sessions using a variety of MD on Tap features in an average session length of less than four minutes. The evaluator, the number and quality of query terms, the total number of citations found for a query, and the use of auto-spellcheck significantly contributed to the probability of query success.

Conclusion: Handheld computers with Internet access are useful tools for healthcare providers to access MEDLINE in real time. MEDLINE citations can answer specific clinical questions when several medical terms are used to form a query. The MD on Tap application is an effective interface to MEDLINE in clinical settings, allowing clinicians to quickly find relevant citations.

■ **J Am Med Inform Assoc.** 2007;14:807–815. DOI 10.1197/jamia.M2424.

Introduction

In the past decade, handheld computers (HHCs) have become important information adjuncts for mobile healthcare providers.[1] Recent advances in HHC hardware and wireless network technology have further facilitated evidence-based practice by supporting real time access to online information systems. However, few published reports address the effectiveness of handheld computers or applications for obtaining online information in clinical environments. This paper describes an evaluation of one such application to seek answers to clinical questions at the point of care using the United States National Library of Medicine (NLM) MEDLINE® database of citations to articles in biomedical journals.

Affiliations of the authors: Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, DHHS (SEH, DD-F, SMH, GL, GRT), Bethesda, MD; John A. Burns School of Medicine (JLJ), University of Hawaii, Honolulu, HI

This study was supported by the Intramural Research Program of the National Institutes of Health, National Library of Medicine, and by the National Institutes of Health, Office of Evaluation award number 05-121-NLM.

Correspondence: Dina Demner-Fushmen, Lister Hill National Center for Biomedical Communications, National Library of Information, 8600 Rockville Pike, Bethesda, MD 20894; e-mail: <ddemner@mail.nih.gov>.

Received for 03/06/07; accepted for publication: 07/14/07.

Background

Handheld Computers as Information Access Tools for Clinicians

Benefits

A systematic review in 2006 shows that the adoption rate of HHCs by physicians is at its highest yet with 46% of US internal medicine physicians and 75% of residents using them daily.[2] In a 2006 study physicians report that use of HHC references results in safer patient care.[1] Clinical and library staff using handheld devices in a clinical setting concluded that the devices provide enough relevant and quickly accessible information to influence clinical decisions and improve patient safety.[3] HHC-based decision support systems improved antibiotic prescribing in critical care,[4] and NSAID prescribing with respect to gastro-intestinal risk factors.[5] Nursing students using a HHC-based drug reference calculated medication dosages with greater accuracy and speed than students using textbooks.[6]

Barriers

The size and mobility that make HHCs attractive to busy clinicians also create barriers to their widespread adoption and general utility. Many physicians find them small and difficult to use.[7] Even residents with HHC experience had

difficulty performing text entry tasks on such devices.[8] Although nurses found MobileNurse, a prototype point of care system for HHCs, generally helpful for checking medical orders at the bedside, they found that it did not display an adequate amount of information on one screen.[9] Physicians and health care organizations are concerned about reliability and security of HHCs.[10] Organizations are also concerned about their ability to provide training, hardware and software IT support, and an appropriate infrastructure.[10]

Opportunities and Challenges of Online Applications

Current common clinical uses of HHCs are for pharmacopeias or medical calculators,[11] applications in which the database or algorithms are small enough to fit within the relatively limited memory capacity of a HHC, thus not requiring network access. These applications require little text entry, relying instead on the point and tap functions at which HHCs excel.[8] As HHCs with wireless network capabilities become common, opportunities and challenges arise for additional clinical applications. A few studies suggest potential and actual use of HHCs to access online clinical information. Physicians who participated in a study of the use of a HHC clinical reference application recommend additional access to guidelines, online medical texts, and article search and retrieval [Rothchild]. Internal medicine residents who participated in a feasibility study of smart phones used to access online medical resources perceived the approach as being useful for patient care, but reported the small keyboard and screen as negative factors.[12] PalmCIS, a handheld application for accessing patient data, also allows users to query MEDLINE via PubMed®, NLM's search engine, through a search tool called infobuttons. A survey of PalmCIS users showed that a connection/retrieval time of 15 seconds on average was perceived as too slow, which discouraged use of the application.[13] This small set of studies suggests that clinicians perceive the potential value of online HHC clinical applications, but the design of such applications must take into account the inherent limitations of the small device.

The MD on Tap System and Application

The MD on Tap research project was initiated to discover optimum techniques to deliver information to mobile healthcare professionals via HHCs.[14] Lacking inhouse access to a group of healthcare providers, our research strategy is to build a system that any healthcare professional with a HHC with Internet access can use, offer free client software for the user, observe usage patterns, and solicit feedback. MEDLINE was selected as the data source for the system because it is freely available, it is supported by a fast and thorough search engine (PubMed), the NLM e-utilities afford an interface for third party software,[15] and our target users know and trust MEDLINE. The extensive metadata associated with MEDLINE citations support the selective and nonlinear access that is common among medical professionals and consistent with good HHC application design.[16] Furthermore, short summaries of search results, which are appropriate for the small HHC screen and available for MEDLINE citations, are found to be as effective for decision making as longer summaries.[17]

MD on Tap system design, shown in Figure 1, is client-server, using established Internet protocols (i.e., http) to communicate between the HHC client application and the intermediate server.[18] The client software is designed to accommodate the small screen, limited text entry options and narrow data bandwidth typically available in HHCs. Currently we offer clients for PDAs and PDA/cell phones running on the Palm® or Windows Mobile® operating systems. The intermediate server receives search terms and conditions from the clients, formats these into search queries for the user-selected search engine (PubMed by default), and returns the results of the search to the clients in a sparse, tagged format. The intermediate server also delivers user-requested citations from MEDLINE to the client.

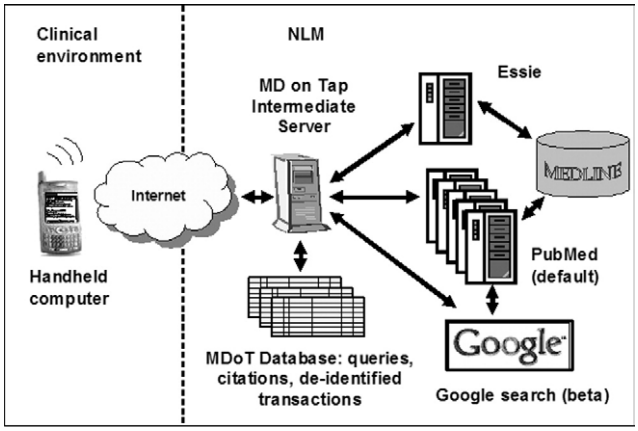

Figure 1. MD on Tap system, showing client, intermediate server and its database, three search engine options, and MEDLINE.

To support data analysis, the intermediate server de-identifies all transactions and stores them in a local database. The analysis of these transactions has revealed several aspects of aggregate user behavior regarding searching MEDLINE with a small device. However, except for the evaluation in the clinical environment discussed later, we do not know the bulk of our individual users or the clinical scenario being addressed by their queries, and therefore do not know if the users are satisfied with their results. Careful manual examination can occasionally infer the clinical question being asked, but with insufficient confidence to determine if the question was answered by the citations retrieved.

By associating search sessions with clinical scenarios and questions, the study described in this paper is designed to learn to what extent users find useful information. Because MEDLINE searching involves both text entry and pointing and tapping, the study also examines the time and actions needed to conduct effective searches.

Methods

Participants and Setting

In conjunction with a clinical elective in Medical Informatics and supervised by clinician faculty members, five residents in internal medicine with training in evidence-based practice conducted explicit evaluation and technology assessment of MD on Tap to answer clinical questions raised at the point of care. These residents/evaluators accompanied medical teams on teaching rounds in teaching hospitals associated with the University Of Hawaii's John A. Burns School of

	Summary of Daily Rounds		Date: MM/DD/YY				
Aprx. Time	Scenario	Clinical Question	EBM Task Cat.	Who Asked	Relevant PMIDs	Comments	Additional Comments
hh:mm am	aged man admitted for sepsis currently has anemia	Does iron supplementation increase or decrease mortality in patient with sepsis in the ICU?	T	A	10966249	CT	Very good article on exactly the topic I was searching for
hh:mm am	elderly woman admitted for sepsis found to have heparin induced thrombocytopenia	How effective is Argatroban in the treatment of heparin induced thrombocytopenia?	T	R	16131539 16125511 16105554	CT	Very good articles even though the drug name was spelled incorrectly (argatroban)

Figure 2. An example of Daily Summary entries (T=Treatment, A=Attending, R=Resident).

Medicine. During rounds, the evaluators used MD on Tap to seek answers in real-time to questions that were raised by any member of the medical team. Because of their broad medical expertise, the evaluators were able to describe and categorize medical scenarios, recognize questions arising in the clinical care environment, formulate search terms related to the questions, search MEDLINE for potential answers to the questions raised, and identify citations relevant to answering the questions.

Teaching rounds were chosen to insure an environment wherein numerous and varied clinical scenarios and questions would be presented. The evaluators rounded with medical teams in the Intensive Care Units and on general medicine wards in two community hospitals. The teams consisted of one attending physician, one chief resident, with up to seven additional residents, and one to six third year medical students. Teams occasionally included a clinical pharmacist and/or an intensivist.

To maximize Internet access from a variety of locations, the evaluators used MD on Tap on a Treo™ 650 PDA/cell phone. While WiFi networks are not yet common in healthcare settings, cell phone coverage with Internet access is widely available. The cell phone data service used by the evaluators was CDMA technology with a basic data rate of less than 100 kilobits per second.

All evaluators had two to five years experience using a PDA, although none had experience with a PDA/cell phone. Prior to the data collection period, each evaluator received introductory training in MEDLINE searching and using MD on Tap from a medical librarian at the medical school. On average, the evaluators spent about 25 minutes training with MD on Tap on the Treo.

The evaluators used a special version of the MD on Tap client that includes a UserID, initialized by the evaluator. That UserId is sent as a tag with each transaction, to allow MD on Tap researchers to associate transactions with evaluator.

Participants' Fieldwork

Each evaluator rounded with teaching teams for about four consecutive weeks. The total period for the five evaluators extended from December 2005 to June 2006. While on rounds, evaluators noted all clinical scenarios and clinical questions, but excluded questions concerning individual patient data such as specific test results. They searched for citations that addressed the questions, and identified those citations judged to be relevant to answering the question.

Evaluators were encouraged to explore all features of the MD on Tap application in order to assess the usefulness of each. On a few occasions, MD on Tap researchers requested that an individual evaluator specifically try one or more of the features. For the most part, however, the evaluators were free to use MD on Tap as they preferred.

Following rounds, evaluators wrote daily summaries consisting of the scenarios, clinical questions, EBM task category for each question, which member of the medical team asked each question, found citations designated by PubMed Identifiers (PMIDs) relevant to answering each question, and comments. Summaries were submitted by email to MD on Tap researchers at NLM. Example records from a daily summary are shown in Figure 2.

Following the data collection period, each evaluator submitted a summative report, a final qualitative and quantitative assessment of the ability to find relevant evidence in MEDLINE using MD on Tap at the point of care. They were asked to distill their experience and discuss MEDLINE searching on a handheld device, features of MD on Tap that were most and least useful for quickly finding relevant citations, and any additional features that could facilitate that process.

Data Collected by the Intermediate Server

The client application sends user requests to the MD on Tap intermediate server for each MEDLINE search query and for each citation selected to be viewed. Details of these transactions, including a timestamp and the UserID, are stored in the MD on Tap database. In addition to queries and citation fetches, the client sends a short message to the intermediate server whenever an evaluator saves a citation to the HHC memory, uses the linkout feature causing the HHC browser to link to the full text of the article offered at a publisher website or through NLM PubMed Central®, or uses the note feature to record comments for later reference.

The intermediate server records the system response time, but not the data transmission time, which is unknown. Our lab tests using the Treo 650 PDA/cell phone hardware and comparable CDMA service measured an average of approximately eight seconds at the HHC to receive either 50 results from a cached query or one cached full citation. This is in contrast to near instantaneous delivery of either data type using MD on Tap on a HHC with WiFi Internet access. We assume that the evaluators experienced approximately the same average eight second delay per transaction as that measured in the lab.

Data Analysis

Sessions

The text recorded by the evaluators in their daily summaries was combined with the transactions in the MD on Tap database to associate each scenario and clinical question with selected PMIDs and with a series of transactions. By

Table 1 ■ Summary of 363 Recorded Clinical Questions and Associated Sessions, with Success rate by Attribute*

Attribute Type/Description	Total Sessions		Successful Sessions	
	Count	% of 363	Count	% of Sessions with that attribute that are successful
Venue†				
ICU	298	82	202	68
Other Bedside	45	12	28	62
Topic Review	20	6	16	80
EBP Task Category†				
Treatment	161	44	113	70
Etiology	125	34	81	65
Diagnosis	55	15	39	71
Prognosis	6	2	3	50
Other	16	4	10	63
Who Asked†				
Attending	127	35	88	69
Resident	206	57	134	65
Intern	27	7	22	81
Student	3	1	2	67
Evaluator‡				
1	55	15	48	91
2	53	15	38	75
3	85	23	60	81
4	109	30	72	65
5	61	17	28	48
Total Questions	363		246	68

*Overall Success rate is 68%.
†not significant.
‡significant at $p < 0.05$.

comparing the approximate time and nature of the question reported in the daily summary with the UserID, time stamps, queries, and citation fetches recorded in the transaction database, we mark the first and last contiguous transaction for that question. We define this set of transactions as the Session associated with the question. One Session may include multiple queries and multiple citation fetches.

Relevant Citations

One author of this paper (SMH), a senior medical librarian and expert indexer, reviewed all of the citations selected by the evaluators and corroborated the relevance of the citation to answering the clinical question as described in the daily summary. Basing her judgment primarily on the abstract, she labeled each citation as A, the abstract contains the major points of the answer and indicates that the paper is useful, B, the abstract covers an aspect of the answer and indicates that the paper will lead to a partial answer, or C, the abstract does not indicate that the paper would be useful and was selected by the evaluator for some other reason. Only those citations that were classified as A in this secondary level of scrutiny were considered Relevant.

Successful Sessions and Queries

For our analysis, a Session is "successful" if at least one of the citations selected by the evaluator as relevant was also classified as Relevant by our expert indexer. An individual query is "successful" if at least one of the citations among the results of the query was Relevant, that citation was viewed by the evaluator during rounds, and it addressed the current clinical question as recorded in the daily summary. A successful Session may include more than one successful query.

Relationships

We examined all unique queries taken from the Sessions, as well as a subset of unique pairs of queries in which the original query was not successful, but a subsequent, different query in the same Session was successful. We analyzed these records for relationships between successful sessions and characteristics of the clinical question, relationships between successful queries and MD on Tap features, and relationships between successful queries and search strategies. Because each transaction record includes a timestamp, we also analyzed elapsed time between certain transactions and for series of transactions.

Statistical Analysis

Logistic regression analysis, Wilcoxon's Signed Ranks test, and Siegel-Tukey test from SAS/SUDAAN version 9.1 were used to test for significant differences. We report differences as significant at the 95% confidence level ($p < 0.05$).

Results

The Clinical Questions and Sessions

Each evaluator accompanied medical teams on teaching rounds of one to two hours per day. Altogether, the five evaluators rounded 77 days, recording 228 observed scenarios and 363 clinical questions. The evaluators found one or more Relevant MEDLINE citations for 246 (68%) of the 363 clinical questions during rounding Sessions. They identified a total of 478 Relevant citations, an average of 1.3 per question or an average of 1.9 per successful Session. Table 1

Table 2 ■ Summary of 363 Session Duration and Activities for Five Evaluators

	Session Duration (sec)†		Session Activities								
			All Transactions*			Queries*			Citation Fetches*		
				Per Session			Per Session			Per Session	
Evaluator	Avg.	s.d.	Total	Avg.	s.d.	Total	Avg.	s.d.	Total	Avg.	s.d.
All	221	210	2404	6.6	4.5	821	2.3	1.8	1365	3.8	3.1
1	289	220	425	7.7	4.7	134	2.4	1.9	265	4.8	3.2
2	354	268	491	9.3	4.8	180	3.4	2.2	230	4.3	3.1
3	139	128	393	4.6	3.0	181	2.1	1.6	183	2.2	1.6
4	164	172	641	5.9	3.7	187	1.7	1.4	435	4.0	2.9
5	264	215	454	7.4	5.8	139	2.3	1.5	252	4.1	4.1

*not significant
†significant at $p < 0.05$

summarizes attributes of the clinical questions and associated Sessions. Analysis by logistic regression indicates that evaluator (i.e., individual characteristics of training, experience, expertise or interest) is the only significant predictor of a successful Session. Stratification by venue or by evaluator did not affect results. If both venue and evaluator are treated as confounding variables, evidence-based task category is a significant predictor of a successful Session, with treatment and diagnosis questions being answered more frequently than those in other categories.

Table 2 shows averages and standard deviations for Session duration and for the primary activities within a Session, for all evaluators and for individual evaluators. Sessions lasted, on average, three minutes 41 seconds, plus the unmeasured amount of time the evaluator spent entering the first query and viewing the results of the final transaction. During that period, evaluators executed, on average, about two queries and viewed about four citations. Other occasional transactions include continued queries (i.e., requesting the next 50 citation summaries), related article links, saving citations to the HHC, link-out to a publisher website, or recording a note for later viewing. When we combine Success rate (Table 1) with Session activities (Table 2), logistic regression indicates session duration is a significant predictor of success, while number or type of transaction are not. However, the Seigel-Tukey test finds no significant difference in session duration among evaluators, although it does find significant differences among evaluators in total transactions, number of queries, and number of citation fetches. Together, these statistics indicate that Session duration does not explain the difference in evaluator Success rate.

Characteristics of Successful Queries—All Queries

Of the 821 MD on Tap queries executed by the evaluators, 803 were a unique combination of terms, features and limits. Of these, 276 (34.4%) were successful. A successful query returns at least one Relevant citation that was fetched for display on the HHC following that query. Attributes of the 803 queries were analyzed to determine which contribute to the probability of success. We analyze three possible contributing factors: features of the MD on Tap application, standard PubMed/MEDLINE limits, and search terms.

MD on Tap Features

Features of MD on Tap are auto-spellcheck, clustered results, and PubMed, Essie, or Google™ search engine options. A single query may employ multiple features. For example, Figure 3 shows the Search tab on a Palm HHC with Essie search engine, EBM clustered output, and auto-spellcheck selected. For a more complete description of the MD on Tap feature set, refer to Hauser et al., 2004.[18] For a comparison of the three search engines in the context of this study refer to Hauser et al., 2006.[19] Table 3 shows the usage frequency of each feature for the 803 unique queries, and the percent of those queries that were successful. We used logistic regression to analyze the relationship between the use of features and success. Auto-spellcheck was used extensively and was found to be the only feature significant in contributing to success. Although evaluators elected to view clustered results rather than listed results for more than a third of their queries, it did not significantly contribute to success. Stratification by evaluator did not affect results.

PubMed/MEDLINE Limits

The *abstract only, English, Human, date range, journal subsets, publication type, and clinical query searches* are standard

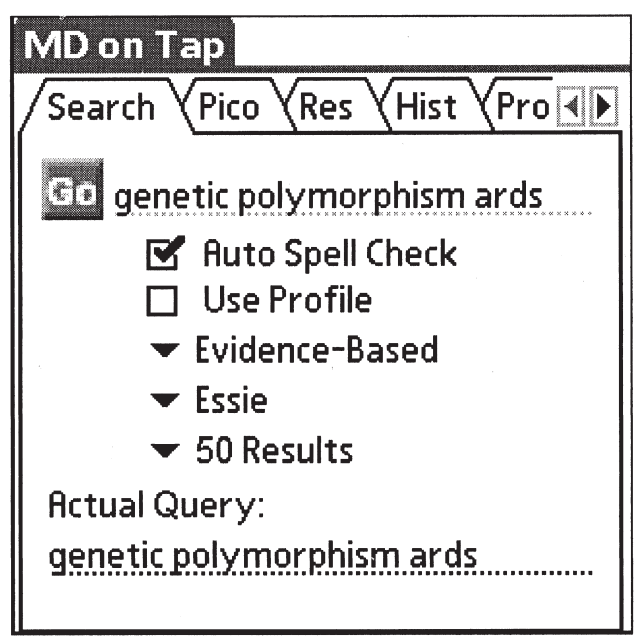

Figure 3. Search tab with Essie, EMB clusters, and auto-spellcheck selected.

Table 3 ■ Use of MD on Tap Features for 803 Queries and Success Rate by Feature, as Compared to the Overall Success Rate of 34.4%

MD on Tap Feature	Frequency Used		% Success of Queries with That Attribute
	Count	% of 803	
Auto-spellcheck*	668	83.2	36.4
PubMed search engine†	629	78.3	35.3
Clustered results†	286	35.6	33.6
Essie search engine†	113	14.1	29.2
Google search engine†	61	7.6	36.1
No features† (other than a search engine selection)	66	8.2	25.8

*significant at $p < 0.05$
†not significant

PubMed/MEDLINE limits that can be selected and applied via the MD on Tap interface. A single query may have several limits. For example, Figure 4 shows the Profile tab with the search limited to articles with abstracts, in English, and about human subjects. Table 4 shows the frequency with which each of these limits was used within the 803 queries, and the percent of those queries that were successful. We again used logistic regression to analyze the relationship between the use of limits and success. None of these limits was significant in contributing to success.

Search Terms

The search terms are unrestricted free text, with the number and quality determined by the evaluator. We analyzed the number of non-stopword terms[1] in the search query and the number of search terms that could be mapped to Medical Subject Headings (MeSH) as identified in the Translation Stack returned from NLM's Entrez E-search e-utility.[20] Logistic regression indicates that both significantly contribute to a successful query. Although the probability of success increases with each MeSH-mappable term up to five, for non-stopword terms, success increases up to three terms and declines somewhat with more terms. Figure 5 shows these relationships for the 803 unique queries in our sample.

[1]Stopword terms are common words that have little or no meaning by themselves, such as *the*, *of*, *on*, and are not indexed for most search engines.

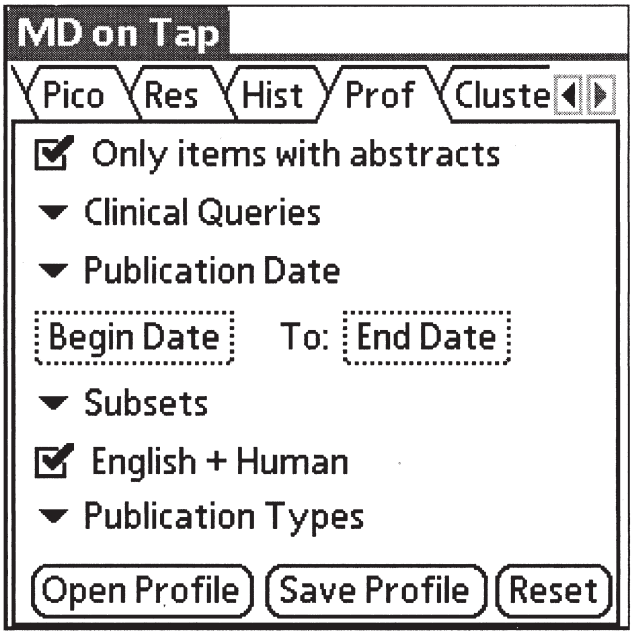

Figure 4. Profile tab with Abstract only and English + Human limits specified.

Characteristics of Successful Queries—Pairs of Queries

From the queries executed during the 363 Sessions, we identified 81 pairs of queries in which the original query in the Session was not successful, but a subsequent, different query in the same session for the same clinical question was successful. We used Wilcoxon's Signed Ranks to test which factors were significantly different in the paired sets. None of the MD on Tap or PubMed features predicted success. Auto-spellcheck was used extensively in both the original and successful queries, so was not a significant distinguishing factor. The use of clustered results was associated with higher success rates, but did not reach significance ($p = 0.1094$.) Significant differences were found for the number of terms, number of MeSH-mappable terms, and the total number of results, i.e., citations found, for the query. We showed the effect of number of terms and MeSH-mappable terms in the previous subsection.

The total number of results is displayed at the bottom of the results tab as shown in Figure 6. The mean and median of the total number of results for the 81 successful queries are 142 and 30, respectively. The mean and median of the total number of results for the 81 original unsuccessful queries are considerably larger at 917 and 92, respectively. However, the notable difference in these two populations is the shape of their distribution. As shown in Figure 7, most (60%) of the successful queries returned a total number of results between six and 100, while only 24% of the unsuccessful queries have a total in this range. The remaining successful queries are about equally divided among those that returned less than six (19%) or more than 100 (21%) total number of results. Of the remaining unsuccessful queries, 28% returned a total less than six and 48% returned a total greater than 100.

The most common change made after an unsuccessful query was to add or modify a search term to narrow the search, for example "diagnosis alpha antitrypsin deficiency" (1619 citations) followed by "management alpha antitrypsin deficiency" (74 citations) which yielded a Relevant article in the first page of results. Evaluators occasionally removed terms

Table 4 ■ Use of Limits for 803 Queries and Success Rate by Limit, as Compared to the Overall Success Rate of 34.4%

PubMed/MEDLINE Limit	Frequency Used		% Success of Queries with That Attribute
	Count	% of 803	
English+Human limit*	273	34	36.6
Abstract only limit*	259	32.3	37.1
Date limit*	117	14.6	42.7
Clinical query hedge*	56	7	37.5
Publication type limit*	25	3.1	32
Journal subset limit*	0	0	NA
No limits*	512	69.2	33.2

*not significant

to broaden the search, for example "spiral ct angiogram arteriogram diagnosis pulmonary embolism" (1 citation) followed by "spiral ct diagnosis pulmonary embolism" (292 citations) which yielded two Relevant articles in the first page of results.

We also examined the transactions that were executed between the unsuccessful queries and the successful queries. For about half (52%) of these pairs, the evaluator executed an additional one to six queries (counting the final query) without looking at any citations. For another 14% of these pairs, the evaluator executed two to seven queries and looked at only one citation. Thus they were able to determine the need for an improved query based on the information available from the Results tab, i.e., article titles and total number of citations, rather than information in the citations. They chose to invest their limited time in a new query with adjusted terms rather than fetching and scanning several citations.

Evaluator Comments

All evaluators said that being able to search MEDLINE at the point of care with a HHC was a positive contribution to evidence-based practice. They thought the primary benefit of this approach is immediate access to the most recently published articles, especially those describing new evidence on a specific topic. The evaluators thought that MEDLINE was an appropriate source of information about specific clinical questions, for example "genetic polymorphisms predisposing patients to ARDS," although less useful as a general topic summary, for example "ARDS etiology, treatment and prognosis." Although interaction with the team was not a planned component of the study, three evaluators indicated that team members often expected answers from them. One evaluator noted on three occasions that the found articles assisted the medical team in patient management.

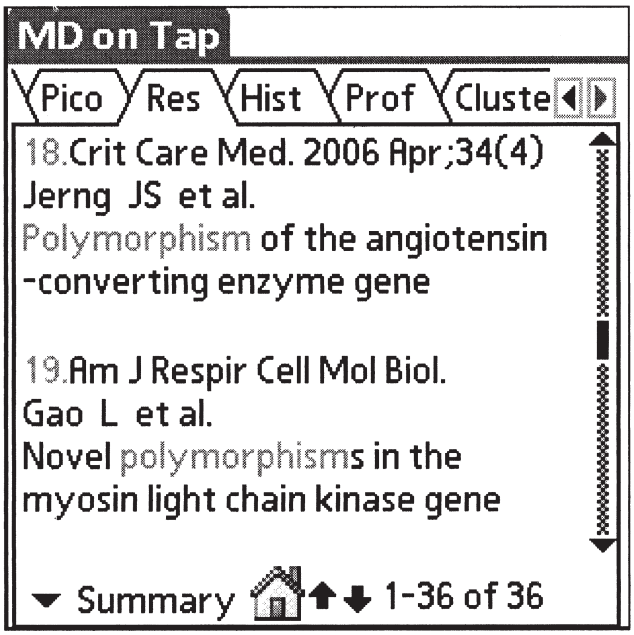

Figure 6. Total number of results (36) shown at the bottom of the Results Tab.

All of the evaluators said that MD on Tap was easy to learn with a short training session conducted by a knowledgeable instructor such as their medical librarian. They also agreed that MD on Tap was generally easy to use. Evaluators told us that within MD on Tap, the article title was the most important field in the citation for deciding whether to retrieve the complete citation, which is consistent with our earlier research.[21] When a citation is then selected and displayed, clinicians are adept at quickly scanning abstract and other text in the citation to assess its value without additional classification aides. There was no consensus on useful and non-useful features of the MD on Tap application.

Most evaluators thought broader adoption of online information retrieval via HHCs will depend on the affordability

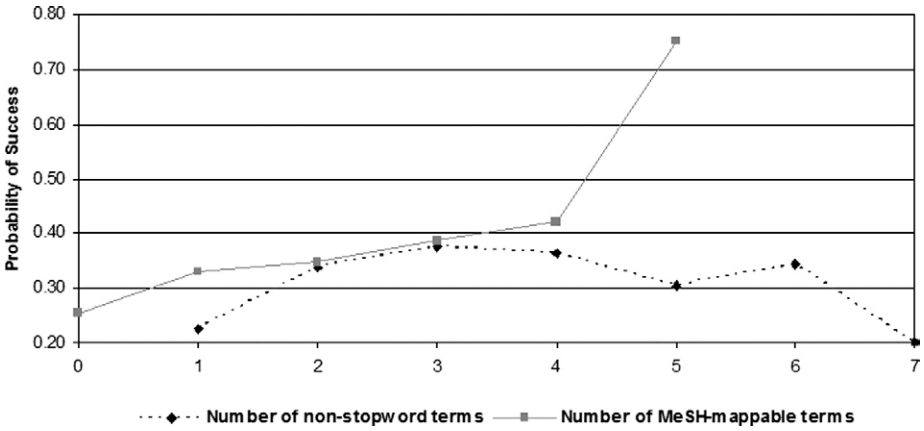

Figure 5. The contribution of number of terms to query success.

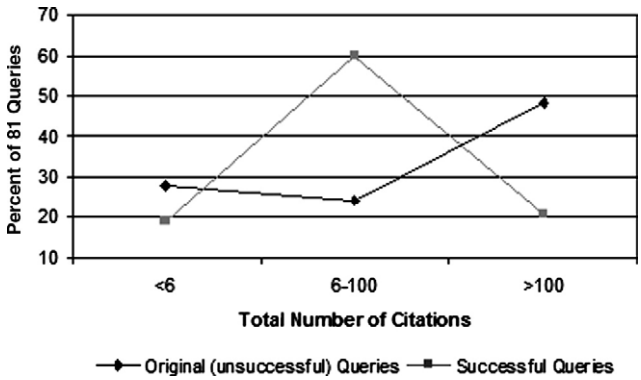

Figure 7. Distribution of 81 query pairs over the total number of results.

of a fast wireless Internet connection at the point of care. Although the small size of the HHC was considered a plus for mobility, it was suggested that display size and data entry may be a limiting factor for many physicians. Evaluators said that seeking answers from MEDLINE using a HHC was useful in many scenarios, but not practical in an ICU. The most common difficulty was the fast pace at which information was presented and questions posed, combined with the relatively slow data rate of the PDA/cell phones. Other difficulties were the many distractions in the ICU and, in their case, their position as observer on the outside of the team.

Discussion

Contributions

We have demonstrated that with a well-designed application, clinical questions can be addressed using a HHC to seek online information at the point of care. This is the first reported analysis of online information searching by clinicians using wireless HHCs in which search actions are automatically recorded and associated with both the questions being addressed and the perceived value of the search results. Prior studies suggest that online search with a wireless HHC can be a useful tool for clinicians, but they do not include quantitative measures of search strategy or search success. Through analysis of the evaluator reports and the automatically recorded data, we are able to identify search strategies and search application features that contribute to fast and successful online searching using a HHC.

Limitations

Although the 363 clinical questions were posed by many clinicians, only five evaluators conducted searches for relevant information. The power of our study is limited by that small number of evaluators, especially in light of their different styles of exploration and information searching. As a result, although we were able to statistically confirm the value of using the automatic spell check feature, other features such as the clinical query hedge or the date limit, which appeared to contribute to search success, were used too infrequently to statistically confirm their value.

Accessing the primary literature to answer questions while on teaching rounds was an artificial situation intended to generate many clinical questions for the evaluation. Evaluator comments indicated that some questions were either too broad or overly specific for MEDLINE to be an appropriate resource. The requirement to address all questions, even when the answer is known, is also artificial, and may have resulted in an overall success rate lower than what would be achieved in actual practice. A clinician who chooses to include MD on Tap as one of a suite of HHC applications would be likely to quickly learn which questions are appropriate for MEDLINE and thus achieve a higher success rate than the 68% reported in this study.

Many clinical questions are only completely answered by the full text of journal articles. Within MD on Tap, if the full text of an article is available, a "link" icon appears at the bottom of the citation screen, as shown in Figure 8. If the user taps the icon, MD on Tap launches the HHC browser with the URL listed in the citation. However, this is rarely practical using a HHC because the URL is usually the journal publisher's web site, which may not be free to the user and is rarely formatted for the HHC. Alternately, the user can save the citation, which includes the URL, on the HHC (the "disk" icon shown in Figure 8) for later reference when a desktop computer is available for more convenient access to the full text.

Implications for Future Research and Development

The demonstrated value of using three or more medical terms for a successful query underlines the importance of both HHC hardware and application user interface. Although text entry is somewhat easier on newer HHCs with small keypads, size remains a barrier to many users. HHC applications for online searching should explore software aids to facilitate text entry, such as word completion and selection from pick lists. The enhanced memory capacity and processing capabilities of newer HHCs would be able to support such tools.

The significant relationship we found between evaluator and success was unexpected and not explained by the data we collected. An information retrieval study in which the

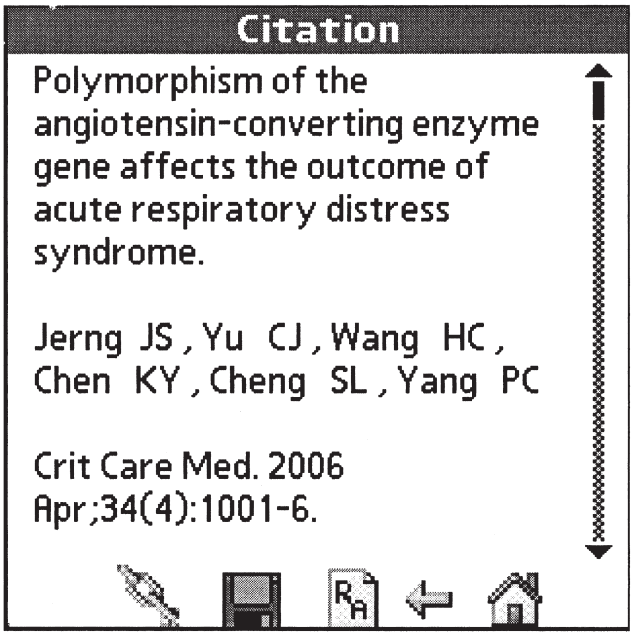

Figure 8. "Full-text link" and "Save-to-memory" icons at the bottom of the Citation screen.

only variable is clinician users with varied backgrounds and training could be useful in determining how to best train clinicians for evidence-base practice, in particular, question formulation, searching for evidence, and critical appraisal skills.

In a 2005 study, 46 clinicians provided up to 20 questions each (698 questions total) and searched online for answers to their own questions using a variety of information resources from desktop computers.[22] Based on their own assessment, the participants found, overall, answers to 73% of their questions with a median search time of approximately 5 minutes. The similarity of success rate and search time to our results suggests a side-by-side comparison of HHC and desktop computer for identical clinical questions and using the same information resources. Such a study could determine the extent to which computer size is a factor in successful information search and retrieval.

Conclusions

Handheld computers with wireless Internet connection are effective platforms for online information delivery to mobile clinicians. Evidence-based practice is encouraged by the quick access to the latest available information, including articles in the primary literature. MEDLINE contains answers to many clinical questions that arise at the point of care, and is a useful adjunct to other clinical information sources. A MEDLINE search application designed expressly for HHCs can be a useful addition to the suite of HHC applications routinely used by mobile clinicians.

To realize the benefits of HHCs, applications developed for mobile clinicians must emphasize system speed and a straightforward user interface. A spell checking option is a valuable aid to mobile clinicians, but must be implemented without incurring extra user actions. MEDLINE searches for answers to clinical questions were most successful when queries consisted of three or more medical terms. To support effective queries, search applications for HHCs must facilitate free text entry on the small computer.

References ■

1. Rothschild JM, Fang E, Liu V, Litvak I, Yoon C, Bates DW. Use and perceived benefits of handheld computer-based clinical references. J Am Med Inform Assoc. 2006;13(6):619–26.
2. Garritty C, El Emam K. Who's using PDAs? Estimates of PDA use by health care providers: a systematic review of surveys. J Med Internet Res. 200612;8(2):e7.
3. Honeybourne C, Sutton S, Ward L. Knowledge in the Palm of your hands: PDAs in the clinical setting. Health Info Libr J. 2006;23(1):51–9.
4. Sintchenko V, Iredell JR, Gilbert GL, Coiera E. Handheld computer-based decision support reduces patient length of stay and antibiotic prescribing in critical care. J Am Med Inform Assoc. 2005;12(4):398–402.
5. Berner ES, Houston TK, Ray MN, et al. Improving ambulatory prescribing safety with a handheld decision support system: a randomized controlled trial. J Am Med Inform Assoc. 2006;13(2):171–9.
6. Greenfield S. Medication error reduction and the use of PDA technology. J Nurs Educ 2007;46(3):127–31.
7. McAlearney AS, Schweikhart SB, Medow MA. Doctors' experience with handheld computers in clinical practice: qualitative study. BMJ 2004;328(7449):1162.
8. Rodriguez NJ, Borges JA, Soler Y, Murillo V, Sands DZ. A usability study of physicians' interactions with PDA and laptop applications to access an electronic patient record system. Proc 17th IEEE Symp Computer-Based Medical Systems 2004;Jun: 153–60.
9. Choi j, Chun J, Lee K, et al. MobileNurse: hand-held information system for point of nursing care. Comput Methods Programs Biomed 2004;74(3):245–54.
10. McAlearney AS, Schweikhart SB, Medow MA. Organizational and physician perspectives about facilitating handheld computer use in clinical practice: results of a cross-site qualitative study. J Am Med Inform Assoc 2005;12(5):568–75.
11. Criswell DF, Parchman ML. Handheld computer use in U.S. family practice residency programs. J Am Med Inform Assoc 2002;9(1):80–6.
12. Leon SA, Fontelo P, Green L, Ackerman M, Liu F. Evidence-based medicine among internal medicine residents in a community hospital program using smart phones. BMC Med Inform Decis Mak 2007;7:5.
13. Chen ES, Mendonca EA, McKnight LK, Stetson PD, Lei J, Cimino JJ. PalmCIS: a wireless handheld application for satisfying clinician information needs. J Am Med Inform Assoc 2004;11(1):19–28.
14. Hauser SE, Demner-Fushman D, Ford GL, Nguyen TQ, Gill M, Thoma GR. MD on Tap: Point of Care Information Delivery—A Report to the Board of Scientific Counselors Technical Report LHNCBC-TR-2005-004, September 2005.
15. U.S. National Institutes of Health, National Library of Medicine. Entrez Programming Utilities. [Online] Available at: http://eutils.ncbi.nlm.nih.gov/entrez/query/static/eutils_help.html. Accessed: February 2007.
16. Swarts J. PDAs in medical settings: the importance of organization in PDA text design. IEEE Trans Prof Comm. 2005;48(2):161–76.
17. Sweeney S, Crestani F. Effective search results summary size and device screen size: Is there a relationship? Inf Process Manage 2006;42:1056–74.
18. Hauser SE, Demner-Fushman D, Ford G, Thoma G. PubMed on Tap: discovering design principles for online information delivery to handheld computers. Medinfo. 2004;11(Pt 2):1430–3.
19. Hauser SE, Demner-Fushman D, Ford GM, Jacobs JL, Thoma GR. Preliminary comparison of three search engines for point of care access to MEDLINE citations. AMIA Annu Symp Proc. 2006:945.
20. U.S. National Institutes of Health, National Library of Medicine. Entrez DTDs. [Online] Available at: http://eutils.ncbi.nlm.nih.gov/entrez/query/DTD/index.html. Accessed February 2007.
21. Demner-Fushman D, Hauser S, Thoma G. The role of title, metadata and abstract in identifying clinically relevant journal articles. AMIA Annu Symp Proc. 2005:191–295
22. Alper BS, White DS, Ge B. Physicians answer more clinical questions and change clinical decisions more often with synthesized evidence: a randomized trial in primary care. Ann Fam Med 2005;3(6):507–13.