Physicians Answer More Clinical Questions and Change Clinical Decisions More Often With Synthesized Evidence: A Randomized Trial in Primary Care

Brian S. Alper, MD, MSPH1,2
David S. White, MD3
Bin Ge, MA4
1DynaMed, Columbia, Mo
2Department of Family and Community Medicine, University of Missouri-Columbia School of Medicine, Columbia, Mo
3A private practice, Columbia, Mo
4Medical Research Office, University of Missouri-Columbia Health Sciences Center, Columbia, Mo

ABSTRACT

PURPOSE Clinicians need evidence in a format that rapidly answers their questions. DynaMed is a database of synthesized evidence. We investigated whether primary care clinicians would answer more clinical questions, change clinical decision making, and alter search time using DynaMed in addition to their usual information sources.

METHODS Fifty-two primary care clinicians naïve to DynaMed searched for answers to 698 of their own clinical questions using the Internet. On a per-question basis, participants were randomized to have access to DynaMed (A) or not (N) in addition to their usual information sources. Outcomes included proportions of questions answered, proportions of questions with answers that changed clinical decision making, and median search times. The statistical approach of per-participant analyses of clinicians who asked questions in both A and N states was decided before data collection.

RESULTS Among 46 clinicians in per-participant analyses, 23 (50%) answered a greater proportion of questions during A than N, and 13 (28.3%) answered more questions during N than A (P = .05). Finding answers that changed clinical decision making occurred more often during A (25 clinicians, 54.3%) than during N (13 clinicians, 28.3%) (P = .01). Search times did not differ significantly. Overall, participants found answers for 263 (75.8%) of 347 A questions and 250 (71.2%) of 351 N questions. Answers changed clinical decision making for 224 (64.6%) of the A questions and 209 (59.5%) of the N questions.

CONCLUSIONS Using DynaMed, primary care clinicians answered more questions and changed clinical decisions more often, without increasing overall search time. Synthesizing results of systematic evidence surveillance is a feasible method for meeting clinical information needs in primary care.

Ann Fam Med 2005;3:507-513. DOI: 10.1370/afm.370.

INTRODUCTION

Evidence-based medicine (EBM), the integration of best research evidence with clinical expertise and patient values,1 promotes clinical decision making based on the most valid research evidence.2-4 But physicians use references they consider fast and likely to provide answers in preference to databases designed to provide research evidence.5,6 Although MEDLINE searches can provide answers for a substantial proportion of family physicians’ questions (43% and 46% in 2 searches conducted by medical librarians), search times are too long (mean times, 27 and 43 minutes, respectively) for point-of-care searching.6,7 Inadequate synthesis of multiple bits of evidence into clinically useful statements is a major obstacle physicians encounter when answering clinical questions.8,9
Solutions are needed to provide clinicians with rapid access to valid clinical knowledge. DynaMed (http://www.DynamicMedical.com) is a potential solution based on 8 specific design and process principles (Table 1). A preliminary study suggested that DynaMed could answer 55% of clinical questions collected from family physicians, and answers could be found more quickly than with electronic textbook collections, but this study did not evaluate actual use during practice. We conducted a randomized trial among primary care clinicians during practice to determine whether use of DynaMed could increase the efficiency of answering clinical questions.

METHODS

Recruitment and Inclusion Criteria
Participants were recruited by 3 methods: (1) online notices to all new registrants (for a 30-day trial of the database) who self-identified as health care professionals (ie, not trainees), (2) postal mailings describing the trial and inclusion criteria sent internationally to about 5,000 members of the Medical Library Association, and (3) similar e-mails sent to about 400 US family medicine residency programs seeking graduates and faculty who met inclusion criteria. It was not practical to document the eligible population because it is unknown how many clinicians were contacted by medical librarians and residency programs.

After registering and requesting participation in the study, clinicians were included if they identified themselves as primary care clinicians who spent at least 20 hours per week in direct patient care and had English fluency, high-speed Internet access, a willingness to enter 20 of their own clinical questions, no previous relationship with DynaMed, and no editorial or financial relationship of their own clinical questions, no previous relationship with DynaMed. Participants provided informed consent through online enrollment and received a 1-year subscription to the database. Previous users of DynaMed were excluded. For participants who listed a specialty other than family medicine, internal medicine, or pediatrics, an author who was blinded to the results (DSW) evaluated their questions for content. If more than one half of their questions were considered specialty focused, all of their questions were excluded.

Randomization and Procedures
Participants were randomized on a per-question basis to be allowed to use DynaMed (A) or not (N). In either case, they could use their usual information sources. No restrictions on sources used were made except that DynaMed use was prohibited for N questions. Questions were assigned in random order uniquely for each participant. As each participant enrolled, a computer-generated random allocation sequence was assigned in blocks of 10 questions (5 A, 5 N) to provide balanced assignments for within-participant analyses. Question order was unaffected by other participants or timing of question entry. To maintain allocation concealment, participants were informed of individual question allocation only after each question was entered. Assignment for individual questions was concealed from investigators until participants had entered data.

Participants were asked to enter clinical questions for searching during practice but were able to enter questions at any time. Online instructions stated, "Please enter the clinical question you wish to search for. Be concise, but add sufficient information so that others can repeat the search. Once you have entered the question, click Start." Data were collected only for questions formally entered as part of the study. Participants entered a clinical question through a free-text dialog box. On clicking the Start button, they were informed whether they could use DynaMed. If allowed (A), a button appeared that would open DynaMed in a new browser window. In either case (A or N), participants had a button that would open a new browser window with the Web site address of the participant's choice entered during study enrollment. Opening a new browser window started the timer (which was not displayed). Participants were instructed to search, read, and think as appropriate for the situation until they believed they either had found an adequate answer or would not find an answer in a reasonable time. On search completion, participants returned to the study window and selected Answer Found or Answer Not Found, automatically stopping the timer.

Participants were presented with their search time, in minutes and sec-

| Table 1. Design and Process Principles for DynaMed |
|---------------------------------|---------------------------------|
| Principle | Description |
| Systematic literature surveillance | Consistent ongoing monitoring of defined information sources considered to have a high yield for primary care |
| Relevance selection filter | Selection of content limited to clinically relevant information |
| Validity selection filter | Selection of new content limited to information that meets or exceeds validity of existing content |
| Concise synopsis | Brief summarization of specific methods and specific outcomes for included content |
| Direct reference citation | Provision of reference support for specific statements and assertions |
| Synthesis of evidence | Revision of outlines and overview statements to reflect all evidence summarized |
| Standardized template | Consistent headings and subheadings for organizing clinical topic outlines |
| Additional content listings | Listing of external reviews and guidelines |
onds, and could change the time recorded if they were interrupted during searching. Participants were asked, “Did the answer make a difference in your clinical decision making?” (if Answer Found) or “Do you think an answer would have changed your clinical decision making?” (if Answer Not Found) and selected Yes or No. All data were recorded by the participants. The content of answers found by participants was not recorded.

**Statistical Analysis**

Based on funding, target enrollment was 50 participants who would each ask 20 questions. We kept the study open as long as possible within the timeline allowed by the grant funding.

For each outcome, we conducted a per-question analysis comparing all A questions with all N questions and per-participant analysis of differences in outcomes within each participant. The results of the per-question analyses are provided as descriptive statistics only because outcomes are not independent of participant, thus violating assumptions used for inferential statistics. This decision was made with statistical consultation before the protocol was submitted for funding, and thus before data collection.

We analyzed the proportion of questions for which an answer was found. For the effect of searching on clinical decision making, we analyzed the proportion of questions for which an answer was found and made a difference, the proportion of questions for which an answer was not found and would have made a difference, and total points assigned for overall impact on decision making. We counted +1 point if an answer made a difference, 0 points if an answer did not change clinical decision making, 0 points if an answer was not found but would not have changed clinical decision making, and –1 point if an answer was not found when an answer would have made a difference. We analyzed median search times for all questions, questions for which answers were found, and questions for which answers were not found. Median times were used because times did not follow a normal distribution.

For the per-participant analysis of proportion of questions answered, we determined the proportion of A questions answered, the proportion of N questions answered, and the difference between these proportions for each participant. This difference was used for statistical analysis. Because this difference was based on paired, nonnormally distributed data, we used the Wilcoxon signed rank test, a nonparametric analog of the paired t test. Per-participant analyses for all other outcomes (eg, proportions of questions for which answers made a difference, median search times) were conducted in the same manner. A normality test was performed on each outcome before Wilcoxon signed rank tests were used, and no outcomes followed a normal distribution. Our null hypothesis for each analysis was that the median difference equals zero and the sum of positive differences equals the sum of negative differences. All analyses were conducted using intention-to-treat principles based on random assignment, regardless of whether participants used the DynaMed database.

To avoid conflicts of interest, the coinvestigator (DSW) and statistician (BG) had full access to all the data and veto-level involvement in inclusion and exclusion decisions for all participants and data. Analysis was conducted by the statistician at an academic center.

This study was approved by the Health Sciences Institutional Review Board of the University of Missouri-Columbia and by the National Science Foundation.

**RESULTS**

**Participants and Clinical Questions**

Eighty-two participants enrolled between January 20, 2004, and April 28, 2004. Sixteen entered no questions. Two previous DynaMed users were excluded. Four participants were excluded because they did not understand instructions, recorded inaccurate data, and did not respond to requests for clarification. The remaining 60 participants entered 780 questions between January 22, 2004, and June 23, 2004. Eight participants (entering 82 questions) were excluded because they were not primary care clinicians and more than 50% of their questions were pertinent to their specialty—cardiology, emergency medicine, general surgery (2 participants), gynecologic surgery, neurology, occupational health and safety, and ophthalmology. Participant and question flow are shown in Figure 1.

Most of the 52 participants were family physicians practicing in the United States who were familiar with computer searching and who highly valued EBM based on self-report (Table 2). Participants used a range of information sources (Table 3). Their questions varied widely in subject area, complexity, and clarity (ie, participants often typed in brief phrases insufficient to determine the actual question). Questions were classified as relating to treatment (45%), diagnosis/evaluation (22%), etiology (4%), adverse effects of treatments or exposures (4%), epidemiology (4%), screening (3%), prognosis (3%), and prevention (2%). Eighty-two questions (12%) were not easily classified.

**Finding Answers to Clinical Questions**

Among the 46 participants who asked at least 1 question in the A state and at least 1 question in the N state (689 questions total), a larger proportion of participants found answers to more questions in the A state (Table 4).

Participants answered 263 (75.8%) of 347 A questions and 250 (71.2%) of 351 N questions (Table 5). Of
the 347 A questions, participants searched DynaMed for 275 (79.3%). Answers were found for 196 (71.3%) of the 275 questions when DynaMed was searched and specifically found with DynaMed for 190 (69.1%).

Effects on Clinical Decision Making
Among the 46 participants who asked at least 1 question each in the A state and the N state, significantly more participants found answers that changed clinical decision making in the A state (Table 4). Of the 263 answers found in the A state, 224 (85.2%, or 64.6% of all questions randomized to A) were noted to make a difference in clinical decision making (Table 5). Of the 250 answers found in the N state, 209 (83.6%, or 59.5% of all questions randomized to N) were noted to make a difference in clinical decision making.

Among the 185 questions for which answers were not found, answers to 150 (81.1%) would have potentially made a difference in clinical decision making (81.0% for A questions, 81.2% for N questions). The per-participant analysis found no significant differences between A and N states.

For the overall impact on clinical decision making, participants generated 156 points from 347 questions in the A state and 127 points from 351 questions in the N state. Per-participant analysis for this measure highly favored the use of DynaMed (Table 4).

Speed of Finding Answers
Time data were invalid (and not recalled by the participant) for 3 questions. Among the remaining 695 questions, the median time for overall searching (whether or not an answer was found) was similar in A and N states (4.95 vs 4.98 minutes) (Table 5).
Median times were similar in A and N states for both successful and unsuccessful searches. Per-participant analyses did not show significant differences (Table 4).

**DISCUSSION**

This randomized trial shows that a database of synthesized evidence helps primary care clinicians answer more of their clinical questions and find more answers that change clinical decision making. Because of the study method, we are unable to determine which features or combination of features (Table 1) is responsible for this outcome.

Demonstrating the advantage of any single resource would be difficult using this study design because clinicians found answers to a high proportion of their questions in the control (N) state. Although previous studies suggested that primary care clinicians fail to find answers to many of their questions, these studies predated frequent Internet use and included questions for which answers were not sought.6-9,14

Search times were the same with and without DynaMed. Finding more answers in the same time suggests that DynaMed increases the efficiency of answering clinical questions.

**Strengths and Weaknesses of the Study**

Allocation concealment with per-question randomization is a major advance for studies evaluating use of databases for answering clinical questions. Trials using per-participant randomization may be confounded by variations between participants and the effects of resource availability on question recognition and selection. Participants in this study could still introduce bias if they applied differing levels of search thoroughness or answer interpretation based on whether they could use DynaMed. Excluding previous DynaMed users and participants who have relationships with medical references avoids known biases, but unrecognized confounders cannot be excluded.
One would expect that adding a resource without limiting use of existing resources would increase the proportion of questions answered. But some participants recorded No Answer Found after searching DynaMed without searching their usual sources. Some participants appeared satisfied that no answers would be found, whereas others appeared to approach the trial as a comparison with usual information sources instead of a comparison with the addition to these sources.

The study was not designed to evaluate degrees of clinical relevance of answers. Participants were asked, “Did the answer make a difference in your clinical decision making?” and were forced to make a dichotomous Yes or No choice. No additional instructions were provided for assisting participants in the interpretation of this item.

The study was not designed to measure the amount of information transmitted per unit time. Times recorded represented total time using information resources and did not separate time spent finding specific answers from time spent reading additional information of interest. Some participants found answers to their questions quickly and then spent more time reading additional information of interest, thus increasing the times recorded.

Participants chose their own questions and used different questions in A and N states. This approach more accurately reflects actual clinical questions than approaches whereby researchers interpret questions asked by clinicians, or clinicians search using questions provided to them. The A and N states were tested using different questions. Per-participant analyses may control for biases between participants but cannot control for within-participant biases introduced by using different questions. Data entry was minimized to facilitate research during clinical practice, so there are insufficient data to determine whether question characteristics differed. With a large sample, randomization should reduce such bias.

**Generalizability for Clinicians**

Clinicians in this study were using familiar search strategies with familiar sources in the control state. The study evaluated the addition of an unfamiliar source. Speed and success rate might increase with practice.

Our findings are likely generalizable to primary care clinicians who use the Internet to answer clinical questions and who value EBM. This study did not include any participants who expressed a low value of EBM. Although 8 specialists were inadvertently entered in the trial (perhaps because of the lack of a definition of primary care on study entry forms), there are insufficient data to analyze the ability of DynaMed to answer questions in subspecialties.

**Future Research**

The principal investigator was the database developer, an approach mandated by the terms of grant support. Multiple procedures (veto-level involvement of coinvestigator, independent statistical assessment, determination of outcome data by participants instead of investigators) were in place to limit the effect of competing

| Measure                                      | With DynaMed No. (%) | Without DynaMed No. (%) |
|----------------------------------------------|----------------------|-------------------------|
| Total number of questions entered            | 347                  | 351                     |
| Questions answered                           | 263 (75.8)           | 250 (71.2)              |
| Questions for which the answer changed decision making | 224 (64.6) | 209 (59.3)              |
| Questions for which participant did not find an answer when the answer would have changed decision making | 68 (19.6) | 82 (23.4)              |
| Overall impact on clinical decision making, points | 156                  | 127                     |
| Median time searching (n = 695 questions), minutes | 4.95                 | 4.98                    |
| Median time to find answers (n = 510 questions), minutes | 4.78                 | 4.89                    |
| Median time for unsuccessful searches (n = 185 questions), minutes | 5.23                 | 5.1                     |
interests. This study’s methods and mechanisms can be adapted for use by independent investigators.

Finding more answers and changing clinical decision making more often would not be beneficial if the answers found were inaccurate. A study using questions from this trial compared the validity of answers found in DynaMed with that of answers found in a combination of 4 other “just-in-time” medical references (see the Supplemental Appendix, available online-only at http://www.annfammed.org/cgi/content/full/3/6/507/DC1). That study found the level of evidence of answers found in DynaMed to meet or exceed that of answers found in the combination for 87% of questions. Given the subjective evaluations inherent in assigning level-of-evidence labels to clinical reference content, however, this study requires confirmation by independent investigators.

Our study design could be adapted by changing allocation parameters to support head-to-head and multiple comparisons. Adding question classification items to data entry and conducting a larger study could allow determination of clinical reference performance specific to question types.

This study was not able to determine which of 8 design and process principles of DynaMed (Table 1) are responsible for providing valid answers to more clinical questions. Determining which factors are important would help inform development of clinical references. We believe that systematic literature surveillance is a critical component. EBM traditionally requires systematic reviews to ensure that relevant evidence is not missed when answering clinical questions. Systematic reviews alone do not have the capacity to meet most information needs during practice, and global efforts expended on systematic reviews are outpaced by research production.15 Systematic literature surveillance provides an information-processing efficiency unattainable with systematic reviews16 and may allow capture of relevant evidence for many more current questions; used alone, however, it may miss earlier research. We believe that in combination with systematic reviews, systematic literature surveillance makes it possible for the primary care knowledge base (ie, the sources primary care clinicians use to access clinical reference information) to reliably represent current best evidence. We would be pleased to be contacted by anyone wishing to advance this research agenda.

To read or post commentaries in response to this article, see it online at http://www.annfammed.org/cgi/content/full/3/6/507.

Key words: Evidence-based medicine; answering clinical questions; databases; systematic literature surveillance; clinical reference; medical decision making; medical informatics

Submitted December 1, 2004; submitted, revised, May 20, 2005; accepted May 23, 2005.

Versions of this report were displayed as posters at the World Organization of Family Doctors (WONCA) and American Academy of Family Physicians Scientific Assembly, October 2004, Orlando, Fla.

Funding support: This material is based on work supported by the National Science Foundation under award 0339256. Any opinion, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. The National Science Foundation was not involved in the design, conduct, analysis, or writing of this study, or the decision to submit this paper for publication. The specific award (Small Business Innovation Research) requires the primary investigator to be employed by the small business conducting the research. To avoid commercial influence on data analysis, the coinvestigator (DSW) and statistician (BG) had full access to all the data and veto-level involvement in inclusion and exclusion decisions for all participants and data. Analysis was conducted by the statistician at an academic center.

Acknowledgments: John E. Hewett, PhD, provided statistical consultation for the study design. Linda S. Zakarian, BA, BS, coordinated research efforts. Sandra Counts, PharmD, Rudolph H. Dresendorfer, PhD, Gary N. Fox, MD, and Daniel C. Vinson, MD, assisted with manuscript revisions.

References

1. Sackett DL, Straus SE, Richardson WS, Rosenberg WB, Haynes RB, eds. Evidence-Based Medicine: How to Practise and Teach EBM. 2nd ed. London, England: Harcourt Publishers Ltd; 2000.

2. Jenicek M, Stachenko S. Evidence-based public health, community medicine, preventive care. Med Sci Monit. 2003;9:SR1-SR7.

3. Phillips B. Towards evidence based medicine for paediatricians. Arch Dis Child. 2003;88:234.

4. Eldredge JD. Evidence-based librarianship: what might we expect in the years ahead? Health Inf Libr J. 2002;19:71-77.

5. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn’t. BMJ. 1996;312:71-72.

6. Ely JW, Osheroff JA, Ebell MH, et al. Analysis of questions asked by family doctors regarding patient care. BMJ. 1999;319:358-361.

7. Ramos K, Linscheid R, Schafer S. Real-time information-seeking behavior of residency physicians. Fam Med. 2003;35:257-260.

8. Chambless ML, Conley J. Answering clinical questions. J Fam Pract. 1996;43:140-144.

9. Gorman PN, Ash J, Wykoff L. Can primary care physicians’ questions be answered using the medical journal literature? Bull Med Libr Assoc. 1994;82:140-146.

10. Ely JW, Osheroff JA, Ebell MH, et al. Obstacles to answering doctors’ questions about patient care with evidence: qualitative study. BMJ. 2000;324:710.

11. Dynamic Medical Information Systems, LLC. DynaMed systematic literature surveillance. 2005. Available at: http://www.dynamicmedical.com/policy. Accessed April 5, 2005.

12. Alper BS, Stevermer JJ, White DS, Ewigman BG. Answering family physicians’ clinical questions using electronic medical databases. J Fam Pract. 2001;50:960-965.

13. Pagano M, Gauvreau K. In: Principles of Biostatistics. Belmont, Calif: Duxbury Press; 1993:276-279.

14. Covell DG, Ulman GC, Manning PR. Information needs in office practice: are they being met? Ann Intern Med. 1985;103:596-599.

15. Mallett S, Clarke M. How many Cochrane reviews are needed to cover existing evidence on the effects of health care interventions? [editorial]. ACP J Club. 2003;139:A11.

16. Alper BS, Hand JA, Elliot SG, et al. How much effort is needed to keep up with the literature relevant for primary care? J Med Libr Assoc. 2004;92:429-437.