

## Research Paper ■

# Do Online Information Retrieval Systems Help Experienced Clinicians Answer Clinical Questions?

JOHANNA I. WESTBROOK, PhD, ENRICO W. COIERA, MBBS, PhD, A. SOPHIE GOSLING, PSYCHD

**Abstract Objective:** To assess the impact of clinicians' use of an online information retrieval system on their performance in answering clinical questions.

**Design:** Pre-/post-intervention experimental design.

**Measurements:** In a computer laboratory, 75 clinicians (26 hospital-based doctors, 18 family practitioners, and 31 clinical nurse consultants) provided 600 answers to eight clinical scenarios before and after the use of an online information retrieval system. We examined the proportion of correct answers pre- and post-intervention, direction of change in answers, and differences between professional groups.

**Results:** System use resulted in a 21% improvement in clinicians' answers, from 29% (95% confidence interval [CI] 25.4–32.6) correct pre- to 50% (95% CI 46.0–54.0) post-system use. In 33% (95% CI 29.1–36.9) answers were changed from incorrect to correct. In 21% (95% CI 17.1–23.9) correct pre-test answers were supported by evidence found using the system, and in 7% (95% CI 4.9–9.1) correct pre-test answers were changed incorrectly. For 40% (35.4–43.6) of scenarios, incorrect pre-test answers were not rectified following system use. Despite significant differences in professional groups' pre-test scores [family practitioners: 41% (95% CI 33.0–49.0), hospital doctors: 35% (95% CI 28.5–41.2), and clinical nurse consultants: 17% (95% CI 12.3–21.7;  $\chi^2 = 29.0$ ,  $df = 2$ ,  $p < 0.01$ )], there was no difference in post-test scores. ( $\chi^2 = 2.6$ ,  $df = 2$ ,  $p = 0.73$ ).

**Conclusions:** The use of an online information retrieval system was associated with a significant improvement in the quality of answers provided by clinicians to typical clinical problems. In a small proportion of cases, use of the system produced errors. While there was variation in the performance of clinical groups when answering questions unaided, performance did not differ significantly following system use. Online information retrieval systems can be an effective tool in improving the accuracy of clinicians' answers to clinical questions.

■ *J Am Med Inform Assoc.* 2005;12:315–321. DOI 10.1197/jamia.M1717.

Online information systems available at the point of care can provide access to up-to-date evidence when a clinical ques-

tion arises. They are potentially one of the most effective interventions to support evidence-based practice in a clinical setting. It has been shown that databases such as MEDLINE contain information relevant to more than half of the clinical questions posed by primary care physicians.<sup>1</sup> When point-of-care online information retrieval systems are available, many clinicians use them<sup>2,3</sup> and report subsequent benefits to decision making and patient care.<sup>2,4,5</sup> However, there have been few empirical studies of professional differences in online resource use<sup>3,6,7</sup> or the extent to which such systems improve the accuracy of clinicians' answers.<sup>8</sup> Limitations of previous studies have included the use of student populations, information retrieval systems that contain only a single resource, usually MEDLINE, and a failure to simulate the time pressures of clinical practice.<sup>9,10</sup>

We conducted an experiment with experienced, practicing clinicians to test the hypothesis that the use of an online information retrieval system improves clinicians' performance in answering clinical questions within a defined time period. We also sought to examine clinician characteristics related to performance. An experimental study design was selected as it allowed us to control a range of variables that would not be possible in a real-world setting and thus enabled us to examine the potential efficacy of an online information system when used by experienced clinicians.

Affiliations of the authors: Centre for Health Informatics, University of New South Wales, Kensington, Australia (JIW, EWC); and Department of Psychology, Royal Holloway College, University of London, Surrey, United Kingdom (ASG).

Dr. Gosling participated in this research while a Senior Research Fellow at the Centre for Health Informatics, University of New South Wales.

Supported by a grant from the National Institute of Clinical Studies.

The authors thank K. Vaughan, K. Lintern, M. Gazarian, S. Stapleton, V. Sinchenko, and B. Booth for sharing their expertise. They are grateful to N. Creswick who assisted with scenario testing, recruitment, and data collection and to M. Wensley, who assisted with recruitment. Thanks to the team responsible for the information retrieval system, M. Walther, K. Nguyen, V. Vickland, and F. Magrabi, the development of which was partly funded by an Australian Research Council Grant.

Correspondence and reprints: Johanna I. Westbrook, PhD, Clinical Centre for Health Informatics, Cliffbrook Campus, University of New South Wales, Kensington, 2052, Australia; e-mail: <j.westbrook@unsw.edu.au>.

Received for publication: 10/11/04; accepted for publication: 01/04/05.

## Methods

### Participants

The sample consisted of 75 clinicians (26 hospital-based doctors [HDs], 18 family practitioners [FPs] [in Australia, FPs spend the majority of their time in community-based medical practice and not treating patients in hospitals], and 31 clinical nurse consultants [CNCs]) who practiced in the state of New South Wales, Australia. HDs from a range of specialties and FPs were recruited over a period of 2 months using invitation letters seeking volunteers sent via mail and e-mail to clinical departments at two major teaching hospitals and via professional organizations representing FPs. Participating HDs were required to have an appointment at a hospital and FPs were required to be currently working in a family practice. Invitations were sent once, and subjects recruited until there was a sufficient number of clinicians to provide the sample size required to test for significant differences between responses obtained with and without the online information system. The study was designed to be able to detect at least a 10% difference in the proportion of correct scenario answers with and without the use of an online information system with 90% power.<sup>11</sup> CNCs were recruited via a CNC list server. CNCs are registered nurses who have at least five years of experience and have completed post-registration qualifications in a specialty area.

### Procedures

Clinicians were asked to participate in a study to investigate health professionals' use of online information retrieval systems, which involved attendance at a university computer laboratory for a two-hour session. Following written informed consent, clinicians were asked to (1) report their years of clinical experience, (2) rate their computer skills on a five-point scale (poor, fair, good, very good, excellent), and (3) rate their frequency of use of online information retrieval systems using a five-point scale (never, once per month, two to three

times per month, once per week, two to six times per week). Clinicians were presented with eight clinical scenarios (see section on development of clinical scenarios below) that posed clinical questions (Table 1). For example, "A mother brings her 15-month-old son who has been seen three times in the past year for glue ear. She has heard that this can lead to learning and developmental problems and thinks her child may need surgery. His hearing is normal. Does current evidence support the need for the insertion of tympanostomy tubes to avoid developmental problems in this child?" The response categories were (1) yes, (2) no, (3) conflicting evidence, (4) don't know. Six scenarios used these four response options, and two scenarios required a one-word answer. Clinicians were able to spend up to 80 minutes answering the scenario questions unaided. No participant exceeded this time, and most took substantially less time for this stage of the study.

The order of scenario presentation was randomized in both stages of the experiment. Directly on completion of the eight questions unaided, clinicians sat alone at a computer workstation and used an online information retrieval system that provided access to six selected sources of evidence including PubMed, MIMS (a pharmaceutical database), Therapeutic Guidelines (an Australian synthesized evidence source focusing on guidelines for therapy [<http://www.tg.com/au/home/index.html>]), *Merck Manual*, Harrison's Textbook and Health Insite (a government-funded consumer-oriented health database [<http://www.healthinsite.gov.au/>]). Five of the six sources presented evidence in a predigested, summarized form with references available for follow-up. The online information retrieval systems used were designed by the software engineering team at the Centre for Health Informatics, University of New South Wales. Clinicians searched by entering key words and could use Boolean operators "AND," "OR," or "NOT." Participants were given a brief written orientation tutorial regarding the system.

Table 1 ■ Clinical Scenarios and Answers

Scenario	Correct Answer
A mother brings her 15-month-old son who has been seen three times in the past year for glue ear. She has heard that this can lead to learning and developmental problems and thinks her child may need surgery. His hearing is normal. Does current evidence support the need for the insertion of tympanostomy tubes to avoid developmental problems in this child?	No, not indicated
What is the best delivery device for effective administration of inhaled medication to a 5-year-old child during a moderately severe acute asthma attack?	Spacer (i.e., holding chamber)
A patient staying in hospital had a myocardial infarction two days ago and is now threatening to sign himself out. You suspect this is due to nicotine withdrawal. The patient wishes to stop smoking and seeks your advice on whether he can start nicotine replacement therapy. Is nicotine replacement therapy appropriate for this patient?	No, use is contraindicated
A 37-year-old woman with infertility comes to see you. She asks about the association between in vitro fertilization (IVF) and breast and cervical cancer. Do women who undergo IVF have a greater risk of breast or cervical cancer than other women of a similar age?	No evidence of increased risk
A woman whose first baby died of sudden infant death syndrome (SIDS) comes to see you. She asks about the risk of her next baby dying of SIDS. Is there an increased risk of SIDS for this woman's next baby?	Yes, there is an increased risk
A 58-year-old woman with long-standing pain of osteoarthritis in knees, hips, and hands asks about the benefits of glucosamine sulfate. Does existing evidence demonstrate that glucosamine has a disease modifying role in osteoarthritis?	Conflicting evidence
A man is bitten by a brown snake and is taken to the hospital emergency department. There is clear evidence of envenoming (poisonous effects of venom). The hospital has run out of brown snake antivenom, so the patient must be given polyvalent snake antivenom (which contains antivenom for all Australian snakes). Should epinephrine be given with the antivenom to prevent anaphylaxis?	Conflicting evidence
What anaerobic microorganism is most commonly found in osteomyelitis associated with diabetic foot?	<i>Peptostreptococcus/Bacteroides</i>

Participants were re-presented with the same eight scenarios and, using the information retrieval system, were asked to find and document the evidence to support their answers to each of the scenarios. Participants were asked to work through the scenarios as they would within a clinical situation and not spend long periods on any one question. Clinicians were informed that they should spend no more than 10 minutes on each question. We did not police this for each scenario but obtained actual search times from the computer logs. Individual clinicians in the study indicated when they had finished, and no participants exceeded the 80 minutes allocated.

A researcher was present throughout the session and gave minimal technical assistance when necessary. Ethics approval for the study was received from the University of New South Wales Human Research Ethics Committee.

### Scenario Development

Eight scenarios of varying complexity were devised by an expert panel of six clinicians (two of whom teach evidence-based medicine within a university medical program), representing three clinical disciplines in consultation with the research team. All scenarios were based on real-life cases and the questions these generated for the clinicians involved. Some scenarios were designed to be within the general knowledge of most clinicians. Others were designed so it was unlikely that the answer would be known before searching. This was to enable both "confirming" and "exploratory" searches to be tested. The scenarios were chosen to be clinically relevant and able to be answered using the knowledge resources available from the information retrieval system.

Members of the expert panel presented possible scenarios with correct answers for consideration. These experts accessed relevant literature and their professional colleagues and drew on their own expert knowledge in determining and verifying the correct answers to the scenario questions. For each selected scenario, the correct answer (either a one-word response or one of three response categories: [1] yes, there is evidence to support, [2] no evidence to support, [3] conflicting evidence), along with sources of evidence to support the correct answer (e.g., a specific journal article or guidelines), were identified. Table 2 provides an example of this information for one of the scenarios. The expert panel, a medical librarian, and the research team undertook searches using the online information retrieval system to validate that the evidence required to answer the scenario questions was available. Through this process, a pool of 27 scenarios was narrowed to eight. Scenarios were rejected if the evidence was not available on the online system or if they were in a highly specialized area.

### Statistical Analyses

Clinicians' written responses to the scenario questions pre- and post-system use were compared. Scenario answers provided before use of the online information system (stage 1) were coded as "correct" according to the expert panel's pre-determined scenario answers and answers provided after system use (stage 2) were coded as correct if the answer was correct AND a relevant evidence source was documented (e.g., the name of a journal article or a therapeutic guideline). Stage 1 answers of "don't know" were classified as incorrect. Stage 2 answers with no documented information source

**Table 2 ■ Example of a Clinical Scenario Profile Documenting the Evidence Base for the Scenario Answer**

#### Scenario

A patient staying in hospital had a myocardial infarction two days ago and is now threatening to sign himself out. You suspect this is due to nicotine withdrawal. The patient wishes to stop smoking and seeks your advice on whether he can start nicotine replacement therapy. Is nicotine replacement therapy appropriate for this patient?

#### Answer

No

#### Possible evidence sources to support scenario answer

Assessment and treatment of smoking

Silagy C, Mant D, Fowler G, Lancaster T. Nicotine replacement therapy for smoking cessation. In: Cochrane Collaboration Library. Issue 4. Oxford: Update Software, 1998.

Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom KO. The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict.* 1991;86:1119-27.

Silagy C, Mant D, Fowler G, Lodge M. Meta-analysis on efficacy of nicotine replacement therapies in smoking cessation. *Lancet.* 1997;343:139-42.

Tang JL, Law M, Wald N. How effective is nicotine replacement therapy in helping people to stop smoking? *BMJ.* 1994;308:21-6.

#### Source: Therapeutic Guidelines

"Assist with smoking cessation"

"Adverse reactions and contraindications": extract

"Nicotine replacement should not be used in pregnancy or while breastfeeding. Nicotine is teratogenic in animals and increases fetal heart rate in humans. It may also have effects on central nervous system maturation."

"Nicotine has the potential to cause dangerous cardiovascular effects in patients with ischaemic heart disease. It is likely that nicotine replacement is less harmful than continued smoking, but in patients with recent myocardial infarction or with severe cardiac arrhythmias, nicotine replacement should not be used. Adverse effects are more likely if patients continue to smoke while they use NRT."

were classified as incorrect or excluded from certain analyses as specified below.

To examine changes in the direction of answers pre- and post-test, scenario answers were categorized using the classification below.

1. Wrong Wrong (WW): Wrong answer before online information retrieval system use and wrong answer after system use [system did not help]
2. Wrong Right (RW): Wrong answer before but right answer after [system helped]
3. Right Wrong (RW): Right answer before but wrong after [system leads to error]
4. Right Right (RR): Right answer before and right after use [system possibly helped to confirm answer]

Clinicians who recorded no post-test evidence to support their answers were excluded from the categorization above.

The number of correct pre- and post-test answers was calculated. The Sign test<sup>12</sup> was used to assess the significance of the direction of the change. Differences between the professional groups were compared using chi-square analyses. Tests for

correlated proportions were undertaken to examine changes in pre- and post-test scores within each professional group and for the sample overall. The McNemar test was applied<sup>12</sup> to examine the direction of changes in pre- and post-test answers. Analysis of variance was used to compare background characteristics of the professional groups. Actual search times were extracted from the computer logs of the online information retrieval system.

## Results

At the beginning of the study 76% ( $N = 57$ , 95% CI 66.3–85.7) of clinicians rated their computer skills as good to excellent and 67% ( $N = 51$ , 95% CI 57.4–78.6) reported using an online information retrieval system once per week or more frequently. There were no significant differences between the professional groups in terms of computer skills or frequency of online information retrieval use; however, FPs had significantly more years of clinical experience (Table 3).

In total, the 75 clinicians provided 600 clinical scenario answers. Neither the reported frequency of online information retrieval use nor computer skills were associated with better performance in the experiment (as measured by number of correct post-intervention answers) (respectively,  $F = 1.25$ ,  $df = 5,71$ ,  $p = 0.30$ ;  $F = 0.22$ ,  $df = 4,71$ ,  $p = 0.93$ ). There was also no correlation between years of clinical experience and test performance ( $r = 0.03$ ,  $p = 0.39$ ). The mean time that clinicians took to search using the information retrieval system per scenario was 6.1 minutes (95% CI 4.5–7.7).

### Does the Use of an Online Information Retrieval System Improve Clinicians' Answers to Clinical Questions?

Pre-test, clinicians correctly answered 29% (95% CI 25.4–32.6) of scenario questions. This improved by 21% post-test to 50% (95% CI 46.0–54.0;  $z = 9.58$ ,  $p < 0.001$ ). We examined changes in the direction of clinicians' answers (Table 4). Clinicians were significantly more likely to change their answers from wrong to right than from right to wrong (McNemar  $\chi^2 = 92.2$ ,  $df = 1$ ,  $p < 0.001$ ). For 21 (3.5%, 95% CI 3.1–3.9) answers, a correct response was recorded pre-test, yet post-test, clinicians recorded no evidence to support their response. In a further 22 scenario answers, an incorrect pre-test answer was recorded and no post-test response was given.

### Does the Effect of Online Information Retrieval System Use Differ by Clinical Group?

There were significant differences between the professional groups in their pre-test scenario scores. FPs had the highest

proportion of correct answers prior to system use and CNCs had the lowest ( $\chi^2 = 29.0$ ,  $df = 2$ ,  $p < 0.01$ ) (Table 5).

Each professional group experienced a significant improvement in test scores following the use of the information retrieval system (HDs:  $z = 4.98$ ,  $p < 0.001$ ; FPs:  $z = 4.01$ ,  $p < 0.001$ ; CNCs:  $z = 7.39$ ,  $p < 0.001$ ). CNCs experienced the greatest level of improvement. For each professional group, there were significantly more changes from wrong to right answers than from right to wrong answers (one-tailed McNemar test: CNCs:  $\chi^2 = 92.2$ ,  $df = 1$ ,  $p < 0.001$ ; HDs:  $\chi^2 = 25.9$ ,  $df = 1$ ,  $p < 0.001$ ; FPs:  $\chi^2 = 15.7$ ,  $df = 1$ ,  $p < 0.01$ ) (Table 6). Following use of the information retrieval system, there was no significant difference between professional groups' proportions of correct answers ( $\chi^2 = 2.6$ ,  $df = 2$ ,  $p = 0.73$ ).

## Discussion

Use of an online information retrieval system resulted in a 21% improvement in clinicians' answers to scenario questions. This supports the view that when presented with a set of questions, access to an information retrieval system can provide evidence to inform and improve clinicians' decision-making processes. It is possible that a proportion of the improvement in post-test responses was due to a "second-look" effect in which subjects changed their mind after reconsidering the case. However, clinicians were still required to find evidence to support their answers.

In 21% of scenarios, clinicians obtained a correct answer unaided and then confirmed this correct answer by producing information using the online system. While it might be argued that the online evidence system produced no benefit in these instances, we have shown in our previous work<sup>13</sup> that clinicians report that this situation results in an increase in their confidence in their original answer. This in turn reduces uncertainty and may translate to more efficient and clearer patient management decisions.

We found in 7% of scenarios that clinicians incorrectly changed their answers following system use, demonstrating that for a small proportion of clinicians use of decision-support systems has a potential to introduce errors. New types of errors have been found to be associated with using decision-support and intelligent monitoring systems.<sup>14–16</sup> One of potential relevance here is *automation bias*. This occurs when individuals have a tendency to over-rely on the computerized system. This may lead to *errors of commission* in which individuals respond to information supplied by the computer, even when it contradicts their existing knowledge.<sup>17</sup> We have very limited knowledge about the incidence of these types of errors associated with computer use in the

Table 3 ■ Sample Characteristics of Clinicians

	Mean No. of Years Experience (SD)	Mean Computer Skills Rating (SD)*	Mean Frequency of Online Evidence Use (SD)†
Clinical nurse consultants ( $N = 31$ )	17 (6.1)	3.1 (0.8)	3.7 (1.2)
Hospital doctors ( $N = 26$ )	13 (8.3)	3.1 (1.0)	4.3 (1.2)
Family practitioners ( $N = 18$ )	22 (8.8)	3.1 (0.8)	3.3 (1.6)
F; df	6.6; 2,72	0.2; 2,71	3.1; 2,72
P	0.002	0.98	0.52

SD = standard deviation.

\*1 = poor to 5 = excellent.

†1 = never; 2 = once per month; 3 = two to three times per month; 4 = once per week; 5 = two to six times per week; 6 = every day.

health system, and it is clearly an area requiring further investigation. This result highlights the need for precise measures for determining the performance of information retrieval systems, similar to those required for the evaluation of other medical interventions. It may also be possible that participants in the study would have arrived at a correct result given more time. However, subjects were not forced to answer individual questions within a specific time interval but had an overall time limit for their searches.

Clinical performance in answering questions, either aided or unaided, will depend on the nature of the questions being asked, and it is likely that a different set of questions would have resulted in different quantitative results. Despite having validated that answers to all scenario questions were available from the information retrieval system, for 7% of scenarios, clinicians recorded no evidence to support their answers. This occurred even though in 50% of these cases, clinicians reported correct pre-test answers and therefore knew what information they were seeking. This result may reflect poor searching skills on the part of clinicians in the study, poor information retrieval system design, a reduced impetus to pursue searching given that the exercise was part of an experiment and not a real clinical situation, or a combination of these factors.

Fifty percent of scenarios were incorrectly answered post-test. The absence of benchmark data from other studies of experienced clinicians' use of online information systems prevents drawing clear conclusions regarding probable reasons for this result. Performance in formulating and answering clinical questions is also likely to depend on contextual information<sup>18,19</sup> not available within the short scenarios used in the

current study. Until further research is conducted in this area, reasons for the failure of the system to influence clinicians' performance to a greater extent remain speculative. For instance, is it a human limit (cognitive), a knowledge limit, a skill limit, or a technology limit?

Our findings regarding the differences in the pre-test results of doctors and nurses contrast with those of Hersh et al.<sup>9,10</sup> They studied 66 students (45 medical and 21 nurse practitioner students) who were given access to MEDLINE to answer a set of five questions and found that pre-test scores of the two groups were both 32%. Nurses failed to improve significantly (35% correct) following searching, while medical students improved to 52%.<sup>9</sup> Nurse practitioner students in that study were equally likely to change their answers from WR as from RW, which the researchers argued was related to the nurses' difficulty in judging the evidence.<sup>20</sup> This finding was not reproduced among our sample of experienced specialist nurses. In a smaller, similar study<sup>10</sup> of 29 students (20 medical and nine nurse practitioner students), medical students had higher scores than nurses both pre- and post-searching, but both groups improved their performance by 33% post-searching. Differences in the characteristics of the subjects, the information retrieval systems, and procedures used make direct comparison with these previous studies difficult.

Our sample of experienced practicing clinicians also demonstrated greater efficiency in locating relevant evidence with an average search time of 6 minutes compared with 30 minutes for medical and nursing students.<sup>9,10,21</sup> The online information retrieval system provided access to a number of summarized evidence sources, which may have also contributed to this result. The design of online information retrieval systems is thus likely to have an impact on their effectiveness in the hands of different clinical groups and improvements in design have the potential to enhance clinical performance beyond that reported here.

Clinical nurse consultants had a low proportion of correct pre-test answers; however, following use of an information retrieval system, their performance improved to a level similar to the post-test scores of HDs and FPs. It is possible that the failure to detect a difference between doctors and nurses' post-test scores is a result of a lack of statistical power in the study. To detect differences of approximately 5% in post-test performance between the groups with reasonable power would have required the completion of several thousand scenarios.<sup>22</sup> This is a limitation noted by other researchers studying clinical problem solving.<sup>23</sup> In essence, our results demonstrate that access to an online information retrieval system improved CNCs' performance to a level similar, if not equal, to that of the medical groups. This may be

**Table 4 ■ Changes in Scenario Answers Pre- and Post-online Information Retrieval System Use (N = 557\*)**

Scenario Responses		% (95% CI)	Total No.
Pre-test	Post-test		
Wrong	Wrong	39.5% (35.4–43.6)	220
Wrong	Right	33.0% (29.1–36.9)	184
Right	Wrong	7.0% (4.9–9.1)	39
Right	Right	20.5% (17.1–23.9)	114
		100%	557

CI = confidence interval.

\*For 43 scenarios, no post-test information sources were recorded, and these cases are excluded from this table, which examines changes in the direction of pre- and post-test responses. These scenario responses were excluded because they did not meet the definition of a correct answer as specified in the Methods section, namely, the provision of an answer and an accompanying information source found using the online information system to support the answer.

**Table 5 ■ Responses to Scenarios by Professional Group (N = 600)**

	No. of Correct Answers (%) (95% CI)		
	Hospital Doctors (N = 208 responses)	Family Practitioners (N = 144 responses)	CNCs (N = 248 responses)
Pre-online evidence use	72 (35%) (28.5–41.2)	59 (41%) (33.0–49.0)	43 (17%) (12.3–21.7)
Post-online evidence use	104 (50%) (43.2–56.8)	79 (55%) (46.9–63.1)	115 (46%) (39.8–52.2)
% Improvement	15	14	29

CI = confidence interval; CNCs = clinical nurse consultants.

Table 6 ■ Changes in Answers by Professional Groups (N = 557)\*

Response Category	Hospital Doctors (N = 187)	Family Practitioners (N = 134)	CNCs (N = 236)
Wrong Wrong	66 (35%)	47 (35%)	107 (45%)
Wrong Right	61 (33%)	34 (25%)	89 (38%)
Right Wrong	17 (9%)	8 (6%)	14 (6%)
Right Right	43 (23%)	45 (34%)	26 (11%)

CNCs = clinical nurse consultants.

\*For 43 scenarios, no post-test information sources were recorded, and these cases are excluded from this table, which examines changes in the direction of pre- and post-test responses. These scenario responses were excluded because they did not meet the definition of a correct answer as specified in the Methods section, namely, the provision of an answer and an accompanying information source found using the online information system to support the answer.

explained by a well-known result in clinical problem solving that competence may be case related and depend on preexisting knowledge.<sup>24</sup> This has led to what was initially a quite controversial realization that "knowledge of content is more critical than mastery of a generic problem-solving process."<sup>23</sup> In other words, having excellent clinical reasoning skills alone is insufficient to compensate for lack of specific clinical knowledge about patient management. In essence, providing all professional groups with equal access to content via an information retrieval system assisted in removing professional differences that may have arisen because of previous experience. Use of summarized evidence sources may also be part of the answer. Should this result be replicated by others, it would provide evidence to support those who believe that, in the future, it may be both possible and necessary to devolve some clinical decisions, currently reserved to medical specialists, to a broader group of health care professionals.

The results of this study significantly build on the work of previous researchers<sup>8-10,25</sup> regarding the effectiveness of online information retrieval systems to answer clinicians' questions. The study did, however, have a number of potential limitations. As far as possible, our experiment simulated clinical practice. The eight scenarios covered simple to complex clinical questions. The scenarios spanned a broad range of clinical areas and were based on real-life questions that clinicians had generated in their practices. However, the nature of conducting an experiment of this type requires that scenarios be reasonably straightforward and thus are somewhat limited in being able to fully represent the range of questions arising from clinical practice. Unlike several previous studies, we allowed clinicians to indicate that the evidence regarding some questions was "conflicting" rather than forcing clinicians to provide yes/no responses.<sup>9,21</sup>

Our clinicians were time restricted and on average spent 6 minutes to search for an answer. While this is significantly shorter than reported in previous studies of students,<sup>10</sup> it may still be outside the time constraints of real clinical practice. The experimental situation also prevented clinicians from seeking out other sources of information, such as colleagues.

Because clinicians volunteered for the study, it is possible that they were atypical of the clinician population of interest.<sup>26</sup> For example, it is possible that they had a greater interest in evi-

dence-based medicine, online information retrieval systems, and a higher level of computer skills than the "average" clinician and that these factors may have inflated the benefits of system use. However, we found no relationship between performance and level of computer skills or online evidence experience. Results from a previous randomized survey of more than 5,500 clinicians also indicate that the clinicians in our study had a similar level of computer skills to that reported by other doctors and nurses in the State of New South Wales.<sup>6,7</sup>

## Conclusion

This is the first study, of which we are aware, that shows that access to multiple clinical information resources online enhances experienced clinicians' performance in accurately answering clinical questions. The results add to our limited evidence base regarding tools that are effective in supporting clinician decision making. Such work lays a foundation to inform the design of information retrieval systems that can meet the realities of the clinical environment.

## References ■

- Gorman P, Ash J, Wykoff L. Can primary care physicians' questions be answered using the medical literature? *Bull Med Lib Assoc.* 1994;82:140-6.
- Westbrook J, Gosling A, Coiera E. Do clinicians use online evidence to support patient care? A study of 55,000 clinicians. *J Am Med Inform Assoc.* 2004;11:113-20.
- Cimino JJ, Jianhua L, Graham M, et al. Use of online resources while using a clinical information system. Paper presented at the American Medical Informatics Association Symposium, 2003.
- Lindberg D, Siegal E, Rapp B, Wallingford K, Wilson S. Use of MEDLINE by physicians for clinical problem solving. *JAMA.* 1993;269:3124-9.
- Haynes R, McKibbon K, Walker C, Ryan N, Fitzgerald D, Ramsden M. Online access to MEDLINE in clinical settings. A study of use and usefulness. *Ann Intern Med.* 1990;112:78-84.
- Gosling A, Westbrook J, Spencer R. Nurses' use of online clinical evidence. *J Adv Nurs.* 2004;47:201-11.
- Westbrook J, Gosling A, Westbrook M. Use of point-of-care online clinical evidence retrieval systems by junior and senior doctors in NSW public hospitals. *Intern Med J.* In press 2005.
- Hersh W, Hickam D. How well do physicians use electronic information retrieval systems?: A framework for investigation and systematic review. *JAMA.* 1998;280:1347-52.
- Hersh W, Crabtree K, Hickam D, et al. Factors associated with success in searching MEDLINE and applying evidence to answer clinical questions. *JAMA.* 2002;9:283-93.
- Hersh W, Crabtree M, Hickam D, Sacherek L, Rose L, Friedman C. Factors associated with successful answering of clinical questions using an information retrieval system. *Bull Med Lib Assoc.* 2000;88:323-31.
- Lwanga S, Lemeshow S. Sample size determination in health studies. Geneva: World Health Organization; 1991.
- Siegal S, Castellan N. Nonparametric statistics for the behavioral sciences. 2nd ed. New York: McGraw-Hill; 1988.
- Westbrook J, Gosling A, Coiera E. The impact of an online evidence system on confidence in decision making in a controlled setting. *Med Decis Making.* In press 2005.
- Coiera E. *Guide to Health Informatics.* 2nd ed. London: Arnold; 2003.
- Skitka L. Does automation bias decision-making? *Int J Hum Comput Stud.* 1999;51:991-1006.
- Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc.* 2004;11:104-12.

17. Macklis R, Meier T, Weinhaus M. Error rates in clinical radiotherapy. *J Clin Oncol*. 1998;16:551-6.
18. Whiting P, Rutjes A, Reitsma J, Glas A, Bossuyt P, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med*. 2004;140:189-202.
19. Seol YH, Kaufman D, Mendonca EA, Cimino J, Johnson S. Scenario-based assessment of physicians' information needs. Paper presented at the 11th World Congress on Medical Informatics, 2004, San Francisco.
20. Hersh W. *Information retrieval: A health and biomedical perspective*. 2nd ed. New York: Springer-Verlag; 2003.
21. Hersh W, Pentecost J, Hickam D. A task-oriented approach to information retrieval evaluation. *J Am Soc Inf Sci*. 1996;47:50-6.
22. Altman D. *Practical statistics for medical research*. London: Chapman and Hall; 1991.
23. Elstein A, Shulman L, Sprafka S. *Medical problem solving: an analysis of clinical reasoning*. Cambridge (MA): Harvard University Press; 1978.
24. Elstein A, Scharz A. Clinical problem solving and diagnostic decision-making: a selective review of the cognitive literature. *BMJ*. 2002;324:729-32.
25. Rose L, Crabtree K, Hersh W. Factors influencing successful use of information retrieval systems by nurse practitioner students. Paper presented at the American Medical Informatics Association Annual Fall Symposium, 1998.
26. Friedman C, Wyatt J. *Evaluation methods in medical informatics*. New York: Springer-Verlag; 1997.