# Cognitive Impact Assessment of Electronic Knowledge Resources: A Mixed Methods Evaluation Study of a Handheld Prototype

**Pierre Pluye MD PhD & Roland M. Grad MD MSc CCFP FAAFP**
**Department of Family Medicine, McGill University**
**517 Pine Avenue West, Montreal, Quebec, Canada, H2W 1S4**

## ABSTRACT

*Rationale: We recently proposed a new method to systematically assess the cognitive impact of knowledge resources on health professionals. Objective: To describe promises and shortcomings of a handheld computer prototype of this method. Background: We developed an impact scale, and combined this scale with a Computerized Ecological Momentary Assessment technique. Method: We conducted a mixed methods evaluation study using a 7-item scale within a questionnaire linked to a commercial knowledge resource. Over two months of Family Medicine training, 17 residents assessed the impact of 1,981 information hits retrieved on handheld computer. From observations, log-reports, archives of hits and interviews, we examined issues associated with hardware, software and the questionnaire. Findings: Fifteen residents found the questionnaire clearly written, and only one pointed to the questionnaire as a major reason for their low level of use of the resource. Residents reported technical problems (e.g. screen trouble) or limitations (e.g. limited tracking function) and socio-technical issues (e.g. software dependency). Conclusion: Lessons from this study suggest improvements to guide future implementation of our method for assessing the cognitive impact of knowledge resources on health professionals.*

## INTRODUCTION

We recently proposed an innovative method for systematically assessing the cognitive impact of knowledge resources on health professionals [1]. This method integrates a real-time self-assessment tool within knowledge resources to evaluate their impact in daily practice. While usual assessment tools measure the acquisition or situational relevance of information (e.g. users' satisfaction), our method evaluates users' attitudes and behaviors associated with the cognitive processing of information, namely how users absorb, understand, interpret and may integrate newly acquired information into practice. The present paper aims to describe promises and shortcomings associated with the implementation of a prototype on handheld computer or Personal Digital Assistant (PDA).

## BACKGROUND

Knowledge resources may support evidence-based practice, and may offer benefits for health professionals at the point-of-care or at the moment-of-need, notably on PDAs [2]. A method to systematically measure the cognitive impact of knowledge resources on health professionals in everyday practice would enhance evaluation of these resources. To develop such method, we tested a 7-item prototype (hereinafter), and critically reviewed the world literature. Our literature review suggested a 10-item scale [3]. Combined with Computerized Ecological Momentary Assessment (CEMA), the use of this impact assessment scale constitutes a new method [1]. CEMA is based on real-time completion of computerized questionnaires, and used in behavioral research to reduce recall bias [4].

The prototype under study was derived from our initial exploratory qualitative research. We explored impacts of a commercial handheld knowledge resource on six family physicians using a case study approach [5]. Findings suggested seven types of cognitive impact at four levels: highly positive (practice improvement, learning and recall), moderately positive (reassurance and confirmation), no impact, and negative impact (frustration).

The prototype combined this 7-item scale and the CEMA technique, and was tested in a cohort of 26 residents. The quantitative results of this cohort study demonstrate the feasibility of real-time systematic assessment of self-reported cognitive impact associated with the use of a knowledge resource [6]. The present paper reports qualitative findings from this cohort study, and specifically aims to examine technical and socio-technical issues. Addressing these issues may enable further implementation of our method, and so pave the way toward a new generation of users' evaluation of electronic knowledge resources.

## METHOD

In a cohort of 26 family medicine residents, we conducted a longitudinal field research study using a mixed methods approach that combined quantitative and qualitative data collection and analysis. In our first phase of recruitment, 20 of 23 first-year residents consented to participate. In the

second phase, six second-year residents were included. Our quantitative assessment of 5,160 questionnaires linked to residents' information hits is published elsewhere (95.9% of opened hits being assessed) [6]. In the present qualitative assessment, we focus on a sub-sample of 1,981 information hits (38.4%) made by 17 first-year residents during a two month block of family practice (three lost to follow-up for medical or pedagogical reasons). The rationale for this focus on a sub-sample of hits was the building of a workable homogeneous set of qualitative data (similar residents in a similar training context).

The intervention incorporated a commercial knowledge resource for primary care on PDA (InfoRetriever 2003 and 2004). This resource allows simultaneous searching of seven databases: an electronic textbook (5 Minute Clinical Consult), the database of Patient-Oriented Evidence that Matters (POEMs), abstracts of Cochrane reviews and guideline summaries, as well as clinical decision and prediction rules, diagnostic test calculators, and history and physical exam calculators. With consent, we used an InfoRetriever tracking function to document each opened item of information as an 'information hit' in a log file on participants' PDA. Log files provided specific data on information hits viewed by the resident, with each hit defined by a title and unique identification number, when the information was opened (date and time stamp), and what search strategy was employed. Each hit consists of a piece or a chunk of information corresponding to an explicit and codified element of medical knowledge with a physical form and a unique identifier. Each hit was linked to the mentioned 7-item impact assessment questionnaire that systematically prompted participants for real-time feedback, namely the perceived cognitive impact of their information hits. In other words, hits constitute our smallest units for data collection and analysis.

One co-author (RG) recruited participants from two McGill family medicine teaching units on Fall 2003 (group 1) and Winter 2004 (group 2). Participants received three hours of training in two sessions, one of which was devoted to using InfoRetriever. The second training session fell at the beginning of the two-month family medicine block rotation during which participants attended an Evidence Based Medicine course. Further InfoRetriever training was offered during the eight weekly course seminars. Thus, participants received InfoRetriever training in a reiterative fashion [7]. Participants completed an impact assessment questionnaire for information hits they retrieved on PDA over the mentioned 2-month period, and were interviewed on their searches for information. Ethics

approval was obtained from the McGill University Institutional Review Board.

Quantitative data collection and analysis: The cognitive impact assessment questionnaire is presented in Figure 1. Participants were instructed to complete the questionnaire by replying 'not applicable' in the event of forgotten hits or hits opened by mistake. Reminders encouraged completion of unanswered questionnaires, as these questionnaires were displayed whenever the PDA was opened. Questionnaire responses were added to an InfoRetriever usage log file on each PDA and transferred to our server via the Internet. Data files were hosted in a centralized database (password-protected server) and descriptive statistics were used to assess the cognitive impact of hits.

**Figure 1. Computerized questionnaire**
*Example: Frustration*



*Example: Positive impact*



Qualitative data collection and analysis: Qualitative data consisted of observations, log-reports, archives and interviews. Observations on participant-researcher interactions were documented (for example an exchange of emails to solve a technical problem). Logs were saved as text files (for example participant_22_log-report.txt). The textual content of information hits with self-reported impact was also archived in text-files (for example participant_22_hit_235.txt). Thus, multiple sources of

evidence allowed us to critically examine interviews, notably the coherence between interviews and the textual content of corresponding hits. In addition, qualitative data permitted the identification of searches for information from series of hits.

A search for information consists of a social action defined by an objective, and is comprised of an information hit or a set of information hits [9]. For example, MD04 searched InfoRetriever on October 13, 2003 to support their diagnostic decision-making regarding a patient with suspected primary embolism. MD04 opened two hits to fulfill this objective: a text-book-like hit entitled "primary embolism", and a rule calculator entitled "primary embolism diagnosis". Regarding the latter, MD04 reported a positive cognitive impact (learning). Searches were identified following a three-step procedure. (1) Prior to interviews, log-reports were analyzed by the interviewer, and same day-hour-topic hits assembled into potential searches. (2) During interviews, potential searches were reviewed and usually confirmed by participants. (3) Post interview, extracts of log-reports, archival data and interviews were systematically assigned to each search using NVivo2 software for qualitative data analysis.

The first author (PP) interviewed all participants. PP, with experience in family medicine and expertise in qualitative research, was unknown to participants. Interviews varied in duration from 15 to 120 minutes, and retrospectively scrutinized the context of InfoRetriever usage, as well as searches for information. At the start, three general questions were asked: (1) did you experience technical problems with your PDA or with InfoRetriever? (2) in general, how difficult was it for you to answer the questionnaires on your PDA? and (3) did the questionnaire discourage your use of InfoRetriever?

Interview transcripts were audio taped, transcribed verbatim and analyzed by the authors. We read and coded data according to a 3-step thematic analysis [8]. "Do not read between the lines" was our coding rule. First, PP assigned extracts of transcripts to themes by going back and forth from textual data to themes (e.g. "freeze"). Subsequently, for each theme, RG read theme-related extracts, changed the assignment of five extracts, refined the label of themes (e.g. "crash or freeze") and proposed a re-organization of themes in accordance with a tree structure (e.g. theme "Software trouble" and sub-theme "crash or freeze"). We reached consensus on changes, refinements and structure in a face-to-face meeting. Second, all extracts that were not assigned to themes by PP at step one were regrouped into a single NVivo report, and independently assigned to existing or new themes (e.g. "tracking function restricted to searches with opened hits"). We reached consensus

on new assignments and new themes during a second face-to-face meeting. Third, PP constructed three matrices to classify extracts by residents (rows) and by themes (columns) using the NVivo "matrix intersection" function (A: Technical problems; B: Software & questionnaire limitations; C: Socio-technical issues). In each matrix, cells provide a "one-click" access to relevant extracts, and so maintain a chain of qualitative evidence between raw data and themes (NVivo dataset available on request). Matrices were then automatically exported into an Excel table to summarize our qualitative findings.

**FINDINGS**

The 1,981 hits acquired by 17 residents refer to 314 searches (an average of 2.5 information hits per day over 47.2 days). The number of hits and searches by resident is presented in Figure 2. This distribution suggests three levels of use of InfoRetriever: *low usage* (less than 10 searches: MD05, MD09 and MD11), *moderate usage* (between 10 and 30 searches: MD01, MD02, MD03, MD04, MD06, MD07, MD08, MD12, MD13, MD14, MD15, MD16 and MD17), and *high usage* (more than 30 searches: MD10). Not surprisingly, the number of hits is highly correlated with the number of searches (Pearson $r = 0.62$, $p = 0.01$). The average number of hits per search for all residents is 6.3. However, MD10 and MD13 suggest two atypical patterns of information-seeking behavior as compared to their colleagues. MD10 retrieved on average 3.1 hits per search that may correspond to highly efficient searching. In contrast, MD13 retrieved on average 11.7 hits per search that might refer to less efficient searching. Indeed, for each search containing at least one hit with a report of cognitive impact, both MD10 and MD13 reported a high level of positive impact (practice improvement or learning).

**Figure 2. Number of hits and searches by resident**
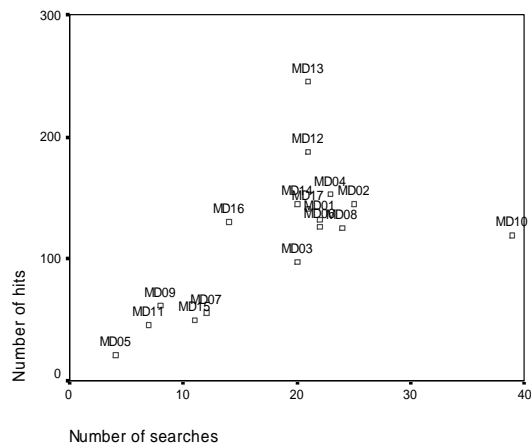
Technical problems or limitations and socio-technical issues reported by residents are presented in Table 1. Overall, 15 residents (88.2%) reported that the questionnaire was clearly written, and as one resident stated "*the questionnaire is generally OK*". However, one resident (5.9%) explicitly pointed to the questionnaire as a major reason for their low level of InfoRetriever usage. As MD09 stated, "the thing that I found frustrating was if you wanted to search something, it would ask you the questions for it, so even if you're just browsing, you're almost taxed for doing that." By contrast, while they moderately used InfoRetriever, five residents reported few questionnaire-related discouragement (MD04: "a little"; MD05: "a little"; MD03: "three or four times"; MD13: "no" then "yes" -maybe-; MD14: "not very much"), and two other residents reported some discouragement (MD08, MD17).

The observation that 14 residents (82.3%) with moderate or high usage of InfoRetriever completed questionnaires when needed, suggests they were able to overcome the following technical problems or limitations and socio-technical issues. (A) Technical problems: Fifteen residents (88.2%) mentioned at least one technical problem. For example, 14 reported that their PDA sometimes froze and needed a soft reset, notably when several applications were open. Of those, two experienced major technical problems that required their PDA to be returned to the manufacturer (screen defect) or a re-install of all software applications (hard reset).

(B) Software and questionnaire limitations: MD07 felt that the InfoRetriever search engine is limited to broad terms, MD15 reported that InfoRetriever provides access to a limited amount of medical information, and MD13 stated that InfoRetriever on PDA may provide empty hits (e.g. a hit that directs the user to a link to a guideline available on Internet). Furthermore, MD06 and MD07 noticed a limitation of the tracking function, which only tracked searches with opened hits. As such, we could not capture searches when users simply read titles of hits but did not open them. Regarding the questionnaire, MD16 regretted there was no option for closing the questionnaire at their convenience, and stated that "if you could just say like wait until later or open it at another time, it would be ideal." Not surprisingly, low users (MD05, MD09, MD11) and the high user (MD10) did not report such limitations.

(C) Socio-technical issues: Residents have established preferences for information sources, and might have been slowed by the questionnaire. Indeed, 11 residents (64.7%) reported that our prototype questionnaire was too time consuming, and as stated by MD13 "just because the switch between screens sometimes was not quick." MD07 suggested an improvement to speed up the process, by implementing a search-level questionnaire that permits the user to exclude irrelevant hits and so reduce the overall number of questionnaires: "cut it down to an assessment of a general search as a whole." Finally, MD10 had the highest level of InfoRetriever use, and reported software dependency: "I am dependent on the machine".

**Table 1. Technical problems or limitations and socio-technical issues reported by residents**

| | *Number of residents reporting problems, limitations or issues* |
|---|---|
| **A. Technical problems** | |
| A.1. Software trouble | |
| A.1.a. Synchronization trouble | *1* |
| A.1.b. Crash or freeze | *14* |
| A.2. Hardware trouble | |
| A.2.a. Defective battery or inadequate capacity | *4* |
| A.2.b. Hard reset | *1* |
| A.2.c. Screen trouble | *2* |
| A.2.d. Trouble opening or closing | *3* |
| ***Residents reporting at least one problem*** | ***15*** |
| | ***88,2%*** |
| **B. Software & questionnaire limitations** | |
| B1. Tracking function restricted to searches with opened hits | *2* |
| B2. Search engine limited to broad terms | *1* |
| B3. Empty hits e.g. link to hit on Internet | *1* |
| B4. Knowledge limitation | *1* |
| B5. No menu option to access favorite hits* | *1* |
| B6. No option for closing questionnaire | *1* |
| ***Residents reporting at least one limitation*** | ***5*** |
| | ***29,4%*** |
| **C. Socio-technical issues** | |
| C1. Questionnaire too time consuming | *11* |
| C2. Software dependency | *1* |
| C3. Potential response pattern | *1* |
| C4. Questionnaire-related discouragement | |
| • Discouragement leading to low use | *1* |
| • Some discouragement (moderate use) | *2* |
| • Little discouragement (moderate use) | *5* |
| C5. Preference for another source of information | *2* |
| C6. Trustworthiness of information | *1* |
| C7. Two questionnaire items perceived as similar (confirmation & reassurance) | *1* |
| ***Residents reporting at least one issue*** | ***14*** |
| | ***82,4%*** |

*This function was available in the 2003 version of InfoRetriever (group 1) but not in version 2004 (group 2).

## DISCUSSION

The present qualitative findings suggest our prototype is promising as it was well-accepted by 16 of 17 residents (94.1%) and did not disrupt their workflow leading to reduced usage of knowledge resource. While completing hit-related questionnaires was perceived as too time-consuming when many hits had been viewed (notably for educational purpose), residents reported that the questionnaire by itself was clear and simple. This problem may be seen as idiosyncratic to the educational context of our study, namely the EBM course that leaded to many information hits, which in turn generated many questionnaires. Indeed, other health professionals may not perceive this problem as they are likely to be more parsimonious in their information-seeking behavior as MD10. Moreover, our findings suggest at least two ways for reducing questionnaire-related burden: (1) decrease the total number of panels, and (2) implement a new search-level panel for deleting hits that are "not applicable".

Surprisingly, 15 of 17 residents overcame technical problems and limitations, as if these problems and limitations were an accepted price to pay for using handheld technology. These problems match usual user complaints about PDAs [10]. Technical advances may permit users to overcome some problems reported by residents in this study. In particular, "new batteries offer longer battery time and memory cards with larger capacity could expand memory storage easily" (p 417).

Our work has at least one limitation. We did not capture residents' searches for information outside InfoRetriever, and so we were not able to put findings into a residency-related information-seeking context. However, our mixed methods evaluation study corresponds to robust designs integrating complementary quantitative and qualitative data collection and analysis [11]. The quantitative part of our research supported the qualitative part and vice-versa. For example, the interview with MD10 provides a convincing qualitative understanding of their high PDA usage, namely software dependency. Such dependency is caricatured as a "palmomental reflex" [12], described in previous qualitative research [13], and fits with our concept of inter-organizational memory [5]. Computers are memory, and in particular clinical information-retrieval technology (e.g. a database of guideline summaries) may be conceived as a collection of electronic memories from multiple organizations.

## CONCLUSION

The present study outlines our experience with a handheld computer prototype of an innovative method for assessing the cognitive impact of knowledge resources used by residents in a training context [1]. Our findings are not generalizable; nevertheless our method may be transferable to other contexts, and can be used to compare knowledge resources from a users' perspective. In addition, information providers may wish to use our method for resource maintenance and when offering credit for point-of-care continuing medical education.

## ACKNOWLEDGEMENTS

## REFERENCES

1. Pluye P, Grad R, Stephenson R & Dunikowski L. A new impact assessment method to evaluate knowledge resources. *AMIA 2005 Symp Proc* 2005:609-13.
2. Cimino JJ & Bakken S. Personal digital educators. *N Engl J Med* 2005;352(9):860-2.
3. Pluye P, Grad R, Dunikowski L & Stephenson R. Impact of clinical information-retrieval technology on physicians: A literature review of quantitative, qualitative and mixed methods studies. *Int J Med Inform* 2005;74(9):745-68.
4. Shiffman S. Real-time self-report of momentary states in the natural environment: Computerized Ecological Momentary Assessment. In: Stone AA, Turkkan JS, Bachrach CA, Jobe JB, Kurtzman HS & Cain VS, editors. *The science of self-report: Implications for research and practice*. Mahwah: L. Erlbaum Associates; 2000. p. 277-96.
5. Pluye P & Grad RM. How information retrieval technology may impact on physician practice: An organizational case study in family medicine. *J Eval Clin Pract* 2004;10(3):413-30.
6. Grad RM, Pluye P, Meng Y, Segal B & Tamblyn R. Assessing the impact of clinical information-retrieval technology in a family practice residency. *J Eval Clin Pract* 2005;11(6):576-86.
7. Grad RM, Meng Y, Bartlett G, Dawes M, Pluye P, Boillat M, et al. Impact of a PDA-assisted Evidence-Based Medicine course on knowledge of common clinical problems. *Fam Med* 2005;37(10):734-40.
8. L'Écuyer R. L'analyse de contenu: Notion et étapes. In: Deslauriers JP, ed. *Les méthodes de la recherche qualitative*. Sillery: Presses de l'Université du Québec; 1987:49-65.
9. Pluye P, Grad RM, Dawes M & Bartlett J. Seven reasons why health professionals search Clinical Information-Retrieval Technology (CIRT): Toward an organizational model. *J Eval Clin Pract*. In press.
10. Lu Y, Xiao Y, Sears A, Jacko J. A review and a framework of handheld computer adoption in healthcare. *Int J Med Inform* 2005;74:409-22.
11. Tashakkori A & Teddlie C. *Handbook of mixed methods in social and behavioral research*. Thousand Oaks: Sage.
12. Crelinsten, GL The intern's palmomental reflex [letter]. N Engl J Med 2004;350(10):1059.
13. Scheck McAlearney A, Schweikhart SB & Medow MA. Doctors' experience with handheld computers in clinical practice: A qualitative study. BMJ. 2004;328(7449):1162-70.