

Decision making in family medicine

Randomized trial of the effects of the InfoClinique and Trip database search engines

Michel Labrecque MD PhD Stéphane Ratté Pierre Frémont MD PhD Michel Cauchon MD Jérôme Ouellet MD
William Hogg MD MSc Jessie McGowan PhD Marie-Pierre Gagnon PhD Merlin Njoya MSc France Légaré MD PhD

Abstract

Objective To compare the ability of users of 2 medical search engines, InfoClinique and the Trip database, to provide correct answers to clinical questions and to explore the perceived effects of the tools on the clinical decision-making process.

Design Randomized trial.

Setting Three family medicine units of the family medicine program of the Faculty of Medicine at Laval University in Quebec city, Que.

Participants Fifteen second-year family medicine residents.

Intervention Residents generated 30 structured questions about therapy or preventive treatment (2 questions per resident) based on clinical encounters. Using an Internet platform designed for the trial, each resident answered 20 of these questions (their own 2, plus 18 of the questions formulated by other residents, selected randomly) before and after searching for information with 1 of the 2 search engines. For each question, 5 residents were randomly assigned to begin their search with InfoClinique and 5 with the Trip database.

Main outcome measures The ability of residents to provide correct answers to clinical questions using the search engines, as determined by third-party evaluation. After answering each question, participants completed a questionnaire to assess their perception of the engine's effect on the decision-making process in clinical practice.

Results Of 300 possible pairs of answers (1 answer before and 1 after the initial search), 254 (85%) were produced by 14 residents. Of these, 132 (52%) and 122 (48%) pairs of answers concerned questions that had been assigned an initial search with InfoClinique and the Trip database, respectively. Both engines produced an important and similar absolute increase in the proportion of correct answers after searching (26% to 62% for InfoClinique, for an increase of 36%; 24% to 63% for the Trip database, for an increase of 39%; $P=.68$). For all 30 clinical questions, at least 1 resident produced the correct answer after searching with either search engine. The mean (SD) time of the initial search for each question was 23.5 (7.6) minutes with InfoClinique and 22.3 (7.8) minutes with the Trip database ($P=.30$). Participants' perceptions of each engine's effect on the decision-making process were very positive and similar for both search engines.

Conclusion Family medicine residents' ability to provide correct answers to clinical questions increased dramatically and similarly with the use of both InfoClinique and the Trip database. These tools have strong potential to increase the quality of medical care.

EDITOR'S KEY POINTS

- Finding the correct answer to clinical questions is an essential first step to evidence-based practice and shared decision making. This randomized trial compared the ability of second-year family medicine residents to find the correct answer to clinical questions using 2 federated medical search engines, InfoClinique and the Trip database. It also aimed to determine users' perceptions of the engines' effects on clinical decision making.
- This trial showed that both InfoClinique and the Trip database provided access to evidence-based clinical information on the benefits and harms of treatment and preventive interventions used in family medicine. Residents also found the 2 search engines similarly useful and easy to use. The choice of one search engine over the other is thus a matter of preference.

This article has been peer reviewed.
Can Fam Physician 2013;59:1084-94

La prise de décision en médecine familiale

Essai randomisé sur l'effet des moteurs de recherche InfoClinique et la base de données Trip

Michel Labrecque MD PhD Stéphane Ratté Pierre Frémont MD PhD Michel Cauchon MD Jérôme Ouellet MD
William Hogg MD MSc Jessie McGowan PhD Marie-Pierre Gagnon PhD Merlin Njoya MSc France Légaré MD PhD

Résumé

Objectif Comparer la capacité de ceux qui utilisent deux moteurs de recherche, InfoClinique et la base de données Trip, à répondre correctement à des questions d'ordre clinique et déterminer ce qu'ils pensent des effets de ces outils sur le processus de prise de décision.

Type d'étude Essai randomisé.

Contexte Trois unités de médecine familiale du programme de médecine familiale de la faculté de médecine de l'Université Laval, à Québec.

Participants Quinze résidents de deuxième année de médecine familiale.

POINTS DE REPÈRE DU RÉDACTEUR

- Répondre correctement à des questions d'ordre clinique représente une première étape essentielle à une pratique basée sur des données probantes et à une prise de décision partagée. Cet essai randomisé comparait la capacité de résidents de deuxième année de médecine familiale de répondre correctement à des questions cliniques en se servant de 2 moteurs de recherche médicaux, InfoClinique et la base de données Trip. L'essai visait aussi à déterminer ce que les utilisateurs pensent de l'effet de ces outils sur la prise de décision clinique.

- Cet essai a montré que InfoClinique et la base de données Trip donnaient tous deux accès à des informations cliniques fondées sur des preuves à propos des avantages et inconvénients des interventions curatives ou préventives utilisées en médecine familiale. Selon les participants, les 2 moteurs de recherche étaient également utiles et faciles à utiliser. Le fait de choisir un moteur de recherche plutôt que l'autre est donc une question de préférence.

Cet article a fait l'objet d'une révision par des pairs.
Can Fam Physician 2013;59:1084-94

Intervention Les résidents ont formulé 30 questions structurées portant sur des traitements curatifs ou préventifs (2 questions par résident) en se basant sur des cas cliniques rencontrés. À l'aide d'une plateforme Internet spécialement créée pour cet essai, chaque résident a répondu à 20 de ces questions (les 2 qu'ils avaient soumises plus 18 autres choisies au hasard parmi celles proposées par les autres résidents, et ce, avant et après avoir cherché l'information dans un des 2 moteurs de recherche. Pour chaque question, 5 résidents ont été assignés de façon aléatoire à commencer leur recherche avec InfoClinique et 5 avec la base de données Trip.

Principaux paramètres à l'étude La capacité des résidents à répondre correctement aux questions cliniques à l'aide des moteurs de recherche, telle que déterminée par l'évaluation d'une tierce partie. Après avoir répondu à chaque question, les participants ont complété un questionnaire destiné à évaluer leur perception de l'effet du moteur de recherche sur le processus de prise de décision en pratique clinique.

Résultats Sur une possibilité de 300 paires de réponses (1 réponse avant et 1 après la recherche initiale) 254 (85%) ont été données par 14 résidents. De ces paires de réponses, 132 (52%) portaient sur des questions pour lesquelles la recherche initiale devait être faite avec InfoClinique et 122 (48%), avec la base de données Trip. Les 2 moteurs de recherche ont entraîné une augmentation importante, et du même ordre en valeur absolue, de la proportion de réponses correctes à la suite de la recherche (de 26% à 62% pour InfoClinique, soit une augmentation de 36%; de 24% à 63% pour la base de données Trip, une augmentation de 39%; $P=0,68$). Pour l'ensemble des 30 questions, au moins un résident a donné toutes les réponses correctes après avoir consulté l'un ou l'autre des moteurs de recherche. Le temps moyen ($M\pm DS$) de la recherche initiale était de $23,5 \pm 7,6$ minutes avec InfoClinique et de $22,3 \pm 7,8$ minutes avec la base de données Trip ($P=,30$). Les participants étaient d'avis que les 2 moteurs de recherche avaient des effets similaires très positifs sur le processus de prise de décision.

Conclusion L'utilisation d'InfoClinique ou de la base de données Trip a entraîné chez des résidents en médecine familiale une augmentation dramatique de leur capacité de répondre correctement à des questions cliniques. Ces outils sont susceptibles de causer une amélioration importante de la qualité des soins médicaux.

Health care professionals' capacity to find the best scientific evidence to answer a clinical question is a key aspect of evidence-based practice.¹ In addition, retrieving information about the benefits and harms of interventions and sharing this information with patients is a key component of shared decision making in clinical practice.²⁻⁴ The Internet now allows clinicians to quickly access a range of pre-appraised evidence synopses, summaries of original research studies, and synthesis facilitating the practice of evidence-based care and shared decision making. According to the 6S hierarchy of pre-appraised evidence,⁵ in the absence of systems providing immediate access to evidence-based clinical information linked to electronic medical records, resources classified as summaries (eg, UpToDate, DynaMed, and Clinical Evidence) should be consulted first to make the practice of evidence-based care most efficient. However, it takes months to years before these tools are updated with the most recent evidence originating from systematic reviews.⁶ This gap might prevent some patients from getting optimal health care. In addition, although these tools can be freely accessed in many academic centres, most individual clinicians have to pay a subscription fee to access information.

Federated medical search engines allow clinicians to search multiple original and pre-appraised sources of evidence at the same time, offering simultaneous access to 5 of the 6S sources of evidence (studies, synopses of studies, synthesis, synopses of synthesis, and summaries). The sites indexed by the search engines might or might not require a subscription fee to be accessed. Numerous federated medical search engines for health professionals are available free of charge on the Web, but very few explicitly prioritize searching sites professing to provide evidence-based clinical information. The Trip database⁷ is probably the most popular of the sites that do. InfoClinique, developed by Laval University in Quebec city, Que, also prioritizes searching evidence-based websites, but allows users to search both French and English websites. Other engines, such as SUMSearch 2, and more recently MacPLUS Federated Search, are limited in the number of the sites they search, either because they only index a few relevant sites (SUMSearch 2) or because users must pay a subscription fee to access many of the sites indexed (MacPLUS).

Only a few studies have assessed the efficacy of different clinical information retrieval tools.⁸⁻¹¹ To the best of our knowledge, no randomized trial has yet compared federated medical search engines or compared such engines to other clinical information retrieval tools. Evidence of the efficacy of such tools would help clinicians wishing to practise according to evidence-based practices and shared decision-making principles to understand the value of these tools and choose the ones that are most appropriate to their contexts.

We conducted a randomized trial to compare the ability of users to find the correct answer to clinical questions, an essential first step to evidence-based practice and shared decision making, searching with 2 federated medical search engines, InfoClinique and the Trip database. We also aimed to determine users' perceptions of the engines' effects on clinical decision making.

METHODS

Participants

The trial was conducted between February and May 2007. We solicited the participation of a convenience sample of all 15 second-year family medicine residents (14 female and 1 male) working in 3 family medicine units of the family medicine program of the Faculty of Medicine of Laval University in Quebec city, Que. All residents agreed to participate and all signed informed consent forms. The trial was approved by the ethics review board of the Saint-François d'Assise Hospital at the Quebec University Hospital Centre.

Pretrial training

In January 2007, before the trial began, all residents participated in a 2-hour training session on how to structure clinical questions in the population, intervention, control, and outcome (PICO) format and to answer questions using InfoClinique and the Trip database. At the same time they were introduced to a website developed to collect data for the project.

Development of clinical questions

Over the following 2 weeks, each of the 15 residents was required to generate 2 PICO format questions about treatments or preventive interventions ("therapy" questions) based on their routine clinical encounters.¹² Questions focused on drug prescriptions—for example, questions about dosage or drug interactions—were excluded. The resident was also required to specify the clinical context in which the question was being asked and to state whether the decision had already been made or had yet to be made. Each question was validated by the member of the research team who supervised the resident who had written the question (M.C., M.L., or P.F.). As they were validated, questions were immediately posted on the study website.

Intervention

The intervention consisted of the use of 2 Internet-based medical search engines, InfoClinique and the Trip database, that specialize in helping users find clinical information. Both tools can be used free of charge. Following the taxonomy used in the 6S model of pre-appraised evidence,⁵ the principal resources indexed in the search engines

are synopses of studies, syntheses (systematic reviews), synopses of syntheses, and summaries (evidence-based clinical practice guidelines and evidence-based textbooks). Both tools also allow easy access to patient information resources and to original studies from the MEDLINE database through a PubMed search using the “clinical queries” strategy.

InfoClinique¹³ was launched by the Department of Family and Emergency Medicine at Laval University in 2003 and was developed from a search engine produced by Coveo Solutions Inc. InfoClinique indexes the content of 74 medical Web resources, in all of which users can access the entire or partial content for free. InfoClinique’s search interface accommodates either English or French; as for the Web resources themselves, 23 are indexed in both French and English, 35 are only indexed in English, and 16 are only indexed in French.¹³ Fifteen resources are specific to health care in the province of Quebec. The websites indexed by InfoClinique are assigned to 1 of the following categories: evidence-based medicine (34 of the 74 websites fall into this category), continuing professional development, complementary and alternative medicine, patient information, professional information, public health, medical images, and electronic textbooks. The user can search all categories at once or can select 1 or more categories. When the user selects all categories, InfoClinique prioritizes results from the evidence-based medicine category when producing the search results. Category filters can also be applied following the search to view results from specific categories.

The Trip database,⁷ produced by Trip Database Ltd of the United Kingdom, was launched in 1997.^{14,15} At the time of the study, the Trip database was searching 489 English medical Web resources at the same time. Access to all pre-appraised evidence websites sites is free with the exception of Clinical Evidence, full texts of articles from the Cochrane Library, and the evidence-based journals from the BMJ Group. New entries are indexed daily, weekly, or monthly, depending on the type of resource. Search results are categorized by type of resource and prioritize evidence-based synopses, systematic reviews, and guidelines. Search results can be limited to a specific type of resource.⁷

Outcome measures

The main outcome measure was the efficacy of each medical search engine, defined as the proportion of correct answers to clinical questions produced by the second-year family medicine residents. Correct answers were determined by consensus. First, on site at Laval University, each resident and his or her supervisor answered the resident’s 2 clinical questions using the information retrieved by all the residents who had searched for an answer to those questions. The answer

produced by the resident and his or her supervisor was then reviewed by 1 of the other 3 supervisors. Second, using the protocol developed for the “just-in-time” project,¹⁶ a trained librarian from the Institute of Population Health at the University of Ottawa in Ontario answered each question. Both answers were then independently reviewed by a research team member at Laval (M.L.) and by another team member at Ottawa (W.H.). These reviewers retained all 30 answers provided by Laval as the criterion standard. The answers generated by the methodology followed at Laval were generally more comprehensive than those provided by the librarian in Ottawa (indeed, the answers produced at Laval included the answers produced in Ottawa).

Two independent assessors (S.R. and J.O.) compared residents’ answers to the criterion standard and classified them as correct or incorrect. Their interrater agreement was good, with κ coefficients of 0.66, 0.61, and 0.78 for answers provided before the search, after the initial search, and after an additional search, respectively. Any discrepancies were resolved by consensus under the supervision of the lead investigator (M.L.). The assessors were blind to the resident, to the medical search engine first used, and to any Web sources consulted after the initial search. Assessors knew, however, whether an answer had been provided before the search, after the initial search, or after the additional search.

Four secondary outcome measures were used to assess residents’ perceptions of how the search engines affected decision making in clinical practice. First, the Comfort with Information for Shared Decision Making scale (CI-SDM) is an 11-item Likert-scale questionnaire assessing respondents’ comfort with knowledge they can use to engage patients in shared decision making with regard to specific clinical questions (5—most comfort to 1—least comfort). The CI-SDM has not been validated, but it was derived from a validated French version of the 16-item Decision Conflict Scale for physicians.¹⁷ Second, the Usefulness of Clinical Information Scale (UCIS) is a 5-item Likert-scale questionnaire based on Shaughnessy and colleagues’ equation (relevance \times validity/work) of the usefulness of clinical information¹⁸ (5—most useful to 1—least useful). Third, the French version of the Impact Assessment Scale (IAS)^{9,19,20} is an ordinal scale of 10 statements about the effect of information retrieved on clinical practice: the respondent checks all statements that apply. Fourth, the Strength of Recommendation Taxonomy (SORT)²¹ measures the respondent’s assessment of the consistency and the quality of the data retrieved as if she or he were to use it to make a recommendation. Using a decisional algorithm, the respondents ranked the data as A, B, C, or not evaluable.

At the end of the trial, residents were asked to complete an online questionnaire about their intention to

use InfoClinique and the Trip database in the future and their view of the usefulness and ease of use of each engine. Questions about residents' intentions were based on the theory of planned behaviour,²² and questions about their perceptions of the engines' usefulness and ease of use were based on the Technology Acceptance Model.²³

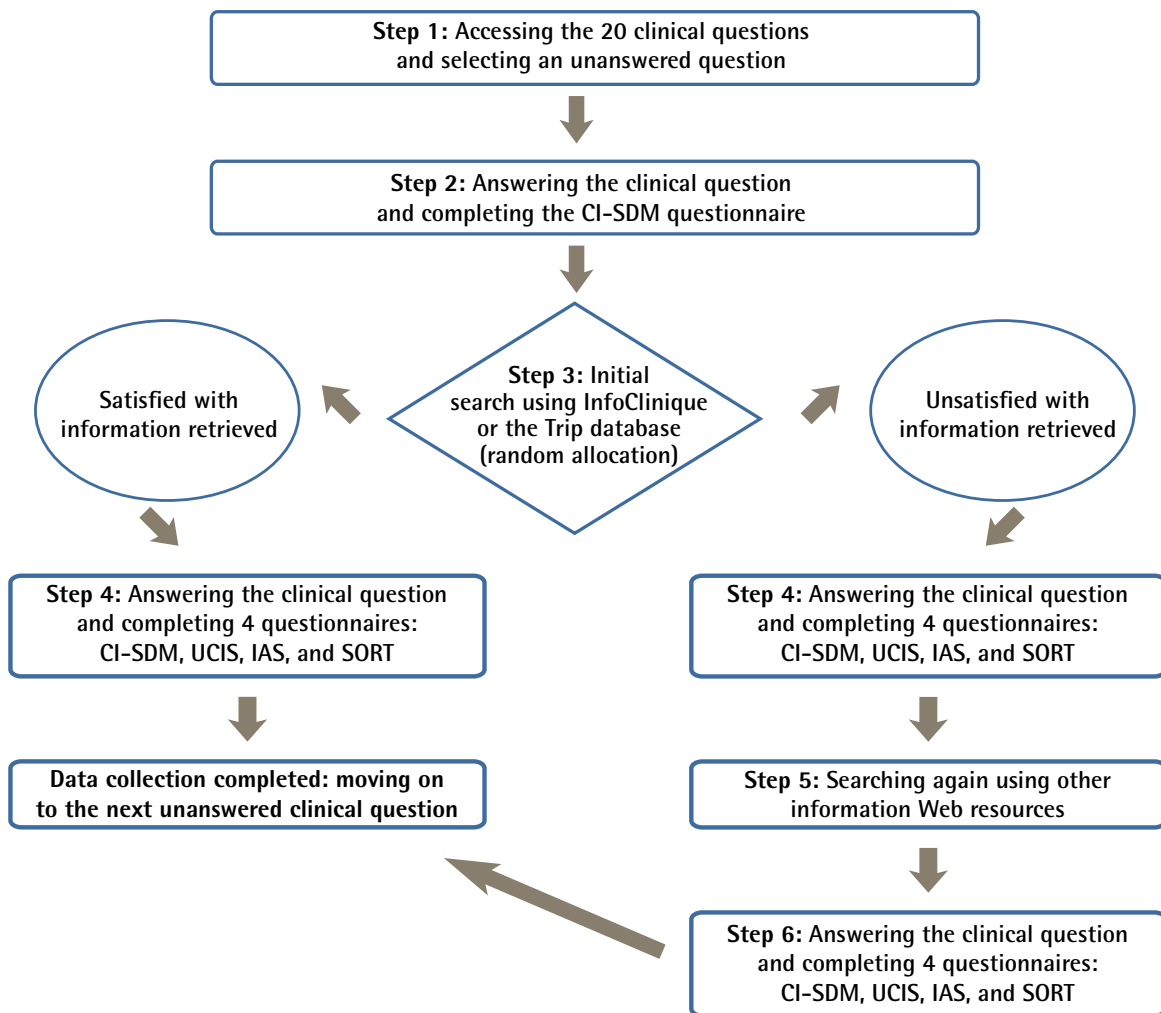
Randomization and data collection

Figure 1 describes the trial and the data collection process using the study website. Residents could access the website with an Internet connection from anywhere and at any time. Upon entering the website with their password, residents were provided with a list of 20 of

the 30 original clinical questions (their own 2 questions, plus 18 questions randomly selected by computer). The questions were displayed in PICO format and ordered according to the time they were posted on the study website. They were instructed to select 1 of their as-yet-unanswered 20 questions (step 1). This number was based on the amount of time allocated to residents' projects at Laval University and the estimated time required for participating in the study (searching for responses and answering clinical questions, completing online questionnaires, and writing a summary of findings for their own 2 questions).

Residents were then required to provide a structured answer to the question, based on their current

Figure 1. Trial and data collection process for each clinical question, using the study website



CI-SDM—Comfort with Information for Shared Decision Making scale, IAS—Impact Assessment Scale, SORT—Strength of Recommendation Taxonomy, UCIS—Usefulness of Clinical Information Scale.

knowledge (step 2). Their answer had to include their best estimates (qualitative and quantitative, if known) of the benefits and harms of each option, including the option of doing nothing. Their answer also had to state the clinical decision that they would make. They were also asked to complete the CI-SDM questionnaire.

When ready to begin searching for the correct answer, residents were randomly prompted by the computer with the home page of either InfoClinique or the Trip database (step 3). For each question, a computer program randomly instructed half of the 10 respondents—the resident who had generated the question and 9 other randomly selected residents—to first perform the search with InfoClinique; the other half were instructed to begin their search with the Trip database. We thus expected residents to perform 300 searches, of which 150 would begin with InfoClinique and 150 would begin with the Trip database. Our random allocation procedure did not seek to assign each participant an equal number of initial searches per search engine: in other words, we accepted that some participants would begin their searches more often with one search engine than with the other. The assignment of individual questions was concealed from investigators.

After searching with either InfoClinique or the Trip database, residents were instructed to select the most relevant and valid information encountered and save it in a temporary file. We recorded search time; up to 30 minutes per question was recommended but was not mandatory. Residents had to stop the timer either when satisfied with their search or when they thought that they had spent sufficient time searching.

After completing the search, residents answered the clinical question, giving their treatment decision based on the information retrieved thus far (step 4). In addition, they had to complete 4 Web-based questionnaires: CI-SDM, UCIS, IAS, and SORT. Residents who were satisfied with the information retrieved by their search could move onto a new question. Residents who were unsatisfied could perform an additional search using any resources available on the Internet, including the alternate medical search engine (step 5). In that case, after completing their subsequent search, they had to answer the clinical question and state their decision in light of the new information retrieved. They also had to complete the 4 questionnaires again (step 6). All the steps for a question were to be completed during the same Web session.

Statistical analysis

Sample size. We based sample size on the time Laval University allowed for residents' projects: namely, 1 half-day per week over 6 months. This time frame permitted a sample of 150 searches with each search engine, enough to capture a difference of 16% (50% vs 66%) in the

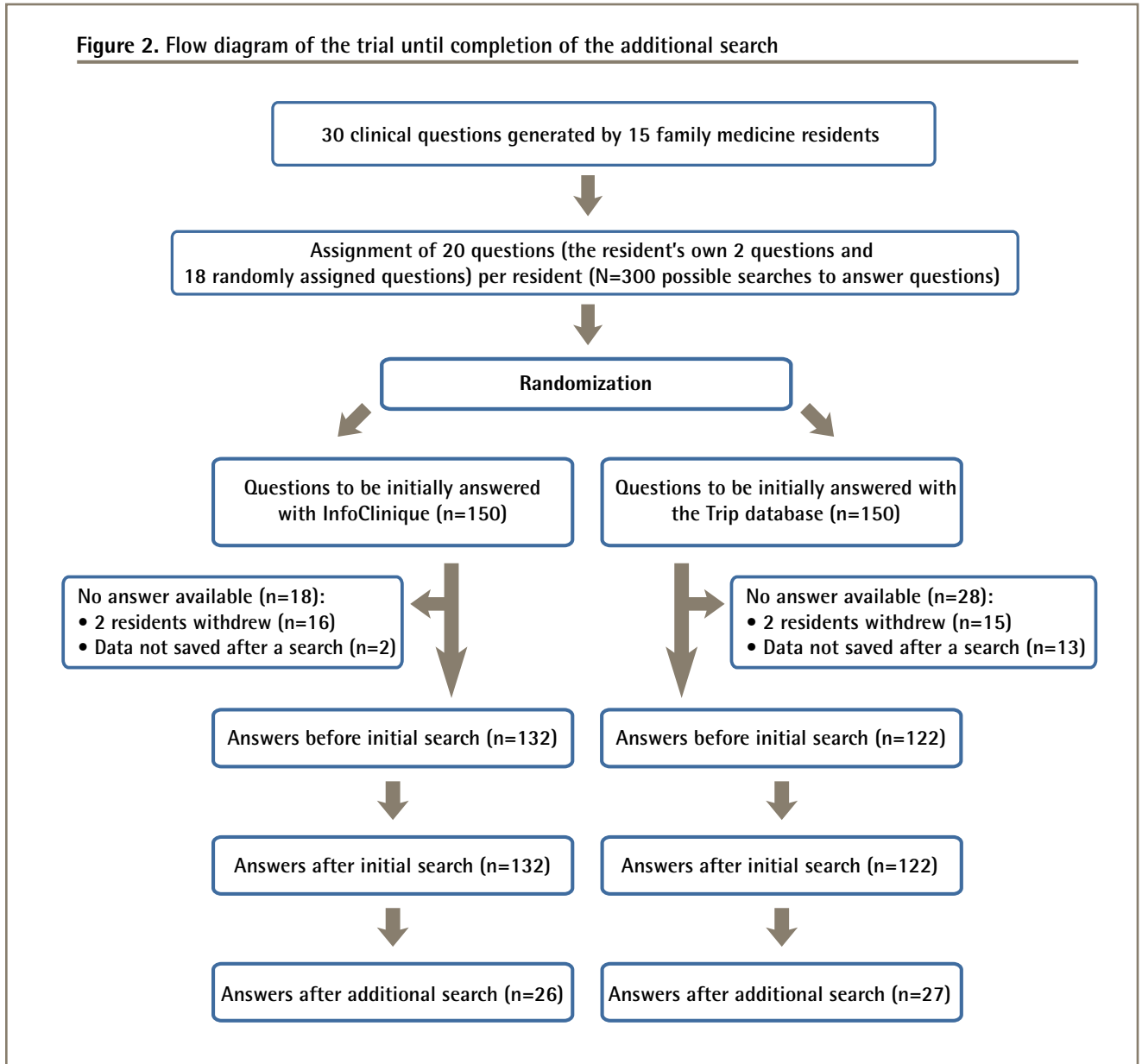
proportion of each study group's correct answers with a power of 80% and α error of .05 in a parallel-group trial design. That said, in our study, this sample size produced greater power because the 15 residents had to randomly use one or the other search engine for each of their 20 searches, acting as their own control. For an exact a priori power calculation, an estimate of the correlation of correct answers resulting from searches with each search engine would have been required. This information was not available before the trial.

Analysis. Our main analyses compared the efficacy of searching with the 2 search engines and the engines' effect on decision making after the initial search (T_1). The primary outcome measure was the proportion of correct answers produced by searching with each engine (efficacy). Results for this measure were compared using the general linear mixed model adjusted for the baseline (T_0) proportions. The general linear mixed model took into account the fact that all residents searched with both search engines and the units of observation were therefore not totally independent. We used the McNemar χ^2 test to evaluate the difference between search engines in the proportion of the 30 questions for which a search produced at least 1 correct answer after the initial search. The difference between the engines' CI-SDM mean scores was assessed using mixed-model analysis of covariance, adjusting for baseline mean scores, and mixed-model analysis of variance to compare UCIS mean scores and mean search times. We used the χ^2 test from the general linear mixed model (polytomous logistic regression) to compare the distribution of IAS and SORT results. Observations reported after the additional search (T_2) represent the cumulative responses of T_1 and T_2 (the last observation carried forward). We limited our analysis of T_2 observations to describing the data: we performed no statistical tests on T_2 data. A paired t test was used to detect differences between search engines regarding residents' intention to use the engines, perceived usefulness, and perceived ease of use. We considered a 2-tailed P value of .05 or less to be statistically significant. All statistical analyses were performed with the SAS statistical package.

RESULTS

Figure 2 presents the trial flow diagram until the completion of the additional search. Of 300 possible pairs of answers (1 answer before searching and 1 answer after the initial search), 254 pairs of answers (85%) were produced by 14 residents (1 resident withdrew before answering any questions). Of these, 132 pairs of answers (52%) concerned questions that had been assigned an initial search with InfoClinique and 122

Figure 2. Flow diagram of the trial until completion of the additional search



pairs of answers (48%) concerned questions that had been assigned an initial search with the Trip database. Twenty-seven of the questions were answered by 8 to 10 residents; the remaining 3 questions were answered by 6 or 7 residents.

Residents' use of InfoClinique or the Trip database for the initial search was quite balanced, with only a few residents using one search engine first considerably more often than they used the other engine first (Table 1). The residents performed 53 additional searches: 26 after an initial search with InfoClinique (20% of all InfoClinique searches) and 27 with the Trip database (22% of all Trip searches). Table 2 presents the residents' location, the residents' level of knowledge, and the level of difficulty of the clinical questions, according to the

search engine used to perform the initial search. The distribution of results was similar for all 3 variables.

Evaluated in terms of residents' capacity to answer the clinical question correctly after performing an initial search, the efficacy of the 2 search engines was very similar (Table 3). There was an important absolute increase in the proportion of right answers after residents searched with either InfoClinique (increase of 36%) or the Trip database (increase of 39%) ($P=.68$). Only 6 answers (2%) that had been correct before the initial search became incorrect after the initial search (5 questions answered with InfoClinique and 1 question answered with the Trip database). These 6 answers were given by 6 different participants for 6 different questions. The increase in the proportion of cumulative correct answers of the initial

Table 1. Number of initial searches performed with InfoClinique and the Trip database by each resident

RESIDENT	NO. OF INITIAL SEARCHES WITH INFOCLINIQUE (N = 132)	NO. OF INITIAL SEARCHES WITH TRIP (N = 122)	DIFFERENCE*
1	12	8	-4
2	9	11	2
3	7	7	0
4	7	13	6
5	11	9	-2
6	10	10	0
7	9	11	2
8	11	9	-2
9	8	10	2
11	7	2	-5
12	7	7	0
13	11	8	-3
14	9	11	2
15	14	6	-8

* χ^2_{13} , $P = .70$ (resident 10 withdrew).

Table 2. Residents' locations, residents' level of knowledge, and level of difficulty of questions, according to the search engine used for the initial search

CHARACTERISTICS	INFOCLINIQUE (N = 132), N (%)	TRIP (N = 122), N (%)
Resident's location		
• FMU 1 (n = 3)	27 (20)	33 (27)
• FMU 2 (n = 5)	45 (34)	40 (33)
• FMU 3 (n = 6)	60 (45)	49 (40)
Resident's level of knowledge before the initial search*		
• High (n = 7)	63 (48)	68 (56)
• Low (n = 7)	69 (52)	54 (44)
Level of difficulty of the clinical question†		
• Easy (n = 6)	27 (20)	24 (20)
• Average (n = 11)	51 (39)	42 (34)
• Difficult (n = 13)	54 (41)	56 (46)

FMU—family medicine unit.

*Based on the proportion of the 20 questions each resident answered correctly before searching (T_0). Residents who answered 15% to 22% and 25% to 43% of the questions correctly were considered to have low and high levels of knowledge, respectively.

†Based on the proportion of residents who correctly answered the question before their initial search (T_0). Questions that were correctly answered by 0% (no correct answer), 10% to 25%, and 88% to 100% of residents were labeled difficult, average, and easy, respectively. No question obtained a score between 26% and 87%.

Table 3. Outcome measures according to the search engine assigned for the initial search: Before searching (T_0), after the initial search (T_1), and after any additional search (T_2).

OUTCOME	INFOCLINIQUE (N = 132)			TRIP DATABASE (N = 122)			P VALUE†
	T_0	T_1	T_2^*	T_0	T_1	T_2^*	
Correct answer, n (%)	34 (26)	82 (62)	88 (67)	29 (24)	77 (63)	82 (67)	.70*
Mean (SD) CI-SDM score [§]	2.6 (0.7)	3.5 (0.8)	3.5 (0.7)	2.6 (0.7)	3.4 (0.8)	3.6 (0.7)	.97
Mean (SD) usefulness of information [§]	NA	3.5 (0.8)	3.6 (0.7)	NA	3.6 (0.8)	3.7 (0.7)	.87 [†]
Assessment of effect, n (%) [#]							.77**
• Strong positive	NA	41 (31)	43 (33)	NA	36 (30)	40 (33)	
• Positive	NA	51 (39)	56 (42)	NA	55 (45)	60 (49)	
• No or negative	NA	40 (30)	33 (25)	NA	31 (25)	22 (18)	
SORT, n (%) ^{††}							.10**
• Level A	NA	36 (34)	35 (33)	NA	51 (48)	52 (49)	
• Level B	NA	50 (47)	52 (49)	NA	36 (34)	37 (35)	
• Level C	NA	20 (19)	19 (18)	NA	19 (18)	17 (16)	

ANCOVA—analysis of covariance, ANOVA—analysis of variance, CI-SDM—Comfort with Information for Shared Decision Making scale, NA—not applicable, SORT—Strength of Recommendation Taxonomy.

*Cumulative values from T_1 and T_2 .

†P values were calculated for the difference between the search engines at T_1 , adjusted for T_0 when possible.

‡General linear mixed model, adjusted for baseline score (score at T_0).

§1—lowest to 5—highest.

||Mixed-model ANCOVA, adjusted for baseline score (score at T_0).

**Mixed-model ANOVA.

‡Strong positive effect—my practice was enhanced; positive effect—I learned something new, I recalled something I had forgotten, I confirmed I was doing the right thing, and I was more confident; no or negative effect—no impact, I am frustrated because there was too much information, I am frustrated because there was nothing useful, I disagree with the information found, and I think the information found could be harmful.

** χ^2 test from general linear mixed model (polynomial logistic regression).

††Level A—consistent and of good quality; level B—inconsistent or of limited quality; level C—based on consensus, usual practice, opinion, disease-oriented evidence, or case series. Residents selected a "not evaluable" option if the criteria did not apply or did not fit the information found. Items for which the user chose "not evaluable" were excluded from the analysis (25 for InfoClinique and 16 for the Trip database).

search and the additional search (T_2) was small: 5% for the InfoClinique group and 4% for the Trip database group. In other words, performing an additional search did not yield an important gain when compared with the yield produced after the initial search.

At least 1 resident produced a correct answer to 26 of the 30 questions (87%) answered after an initial search with InfoClinique and to 28 of the same 30 questions (93%) answered after an initial search with the Trip database ($P = .68$). The 4 questions incorrectly answered with InfoClinique were different from the 2 questions incorrectly answered with the Trip database. This signifies that it was possible to answer all the questions correctly using one or the other medical search engine. All correct answers were retrieved from a free-access website indexed in either one of the search engines.

As measured with the CI-SDM, the UCIS, and the IAS, the residents' perception of the 2 search engines' effect on decision making in clinical practice was similar (Table 3). For the SORT scale, the difference was not statistically significant but the data suggest the possibility of a difference between engines, with more A-level evidence retrieved with the Trip database.

Only the CI-SDM was completed before and after the search. The increase in residents' confidence that their knowledge was adequate for engaging patients in shared decision making regarding a clinical question was identical for the 2 engines. The magnitude of the increase, calculated on the basis of the combined 254 answers, was 0.83 (95% CI 0.70 to 0.95). This corresponds to an effect size of 1.2. An effect size of this magnitude is considered large.²⁴ The mean (SD) time of the initial search for each question was 23.5 (7.6) minutes with InfoClinique and 22.3 (7.8) minutes with the Trip database ($P = .30$).

The mean scores of residents' intentions to use the engines, their perceptions of the engines' ease of use, and their perceptions of the engines' usefulness were high and similar for the 2 engines. The 0.2-point difference in the score of perceived usefulness in favour of InfoClinique, although statistically significant, corresponds to an effect size of 0.2 and thus is not clinically significant (Table 4).²⁴

DISCUSSION

To the best of our knowledge, this is the first randomized trial to compare the effects of 2 federated medical search engines on clinical decision making. Our trial shows that InfoClinique and Trip database searches were similarly effective at helping family medicine residents find the correct answers to clinical questions posed by themselves or by colleagues. The answers found with the 2 engines did not differ in their effect

Table 4. Participants' intention to use the search engines and their perception of the search engines' ease of use and usefulness: $N = 14$.

VARIABLE	INFOCLINIQUE	TRIP DATABASE	P VALUE*
Mean (SD) score [†] of intention to use	6.0 (0.9)	5.6 (1.1)	.16
Mean (SD) score [†] of perceived ease of use	6.0 (0.8)	5.6 (1.3)	.29
Mean (SD) score [†] of perceived usefulness	6.0 (1.0)	5.8 (1.1)	.04

*Paired *t* test.

[†]1—lowest to 7—highest.

on clinical decision making. In addition, we found that searching with either medical search engine greatly improved users' capacity to answer clinical questions correctly and greatly increased their comfort with the knowledge they acquired to engage in shared decision making with patients.

Comparison with previous work

We located only 2 randomized trials comparing the efficacy of searching with different clinical information retrieval tools.^{10,11} In the first,¹¹ searching MEDLINE before searching a selection of evidence-based health care resources was compared with searching the same evidence-based health care resources before searching MEDLINE. In this trial, the proportion of answers assessed by participants as satisfactory was similar in the 2 scenarios. However, only searching MEDLINE produced a higher proportion of satisfactory answers than only searching the evidence-based health care resources did (81% vs 65%). In the second trial,¹⁰ DynaMed, a commercial electronic textbook (classified as a summary according to the 6S hierarchy of pre-appraised evidence⁵), was compared with usual sources of information. The proportion of answers assessed by participants as adequate was similar in the 2 scenarios, with 73% of answers assessed as adequate overall. These trials were marred by 2 methodologic flaws: the failure to record participants' answers before searching, and the decision to use respondents' own determination of a correct or incorrect answer as the outcome measure. Both were avoided in our study.

Three studies have observed an improvement in the proportion of questions answered correctly before and after searching with different clinical information retrieval tools.²⁵⁻²⁷ These studies reported an improvement of 32% among senior medical and nursing students using MEDLINE,²⁵ an improvement of 20% among senior medical students using MEDLINE,²⁶ and an improvement of 27% among physicians and clinical nurses using various tools (MEDLINE, Merck Manual, Harrison's Online,

and others).²⁷ Differences in the clinical question type, in the population of the study, and in the type of clinical information retrieval tools might explain the differences among studies, including ours. It has been observed that a greater global improvement can be expected when physicians pick the clinical question for which they wish to search for an answer, possibly because they believe that a definitive answer exists or for another reason.²⁸ In our study, this possibility was highly attenuated because apart from their own 2 clinical questions, participants did not select the questions they answered.

An additional strength of our study is its design: a randomized trial with residents using both search engines in random order. In addition to minimizing risk of bias, this design increased the power with the sample size available. As the correlation of correct answers resulting from searches with each search engines was 0.45, with 132 and 122 searches with InfoClinique and the Trip database, respectively, our study had a power of 80% to find a difference of at least 14% in the proportion of correct answers between the 2 search engines.

Limitations

Our study also had limitations. First, the questions we studied were limited to therapy- and prevention-related matters of interest to second-year family medicine residents, structured in PICO format. It is likely that we can generalize our findings to primary care clinicians who use medical search engines to answer therapy- and prevention-related clinical questions, but we cannot draw conclusions as to the efficacy of searching with the engines and the effect on clinical decision making for questions addressing prognosis, pathogenesis, or diagnostic procedures.

Second, it is possible that our results would have differed had we used unstructured clinical questions. However, framing questions using the PICO format is recommended as the first step to an effective search¹ and should be applied to all searches for clinical information.

Third, participants were second-year residents in family medicine. This might be a limitation to external validity. Virtually all our residents were fully skilled in using electronic communication and information tools and were at the peak of their medical knowledge—the study was conducted during the months just preceding their licensure final examination. Other type of users might be less effective in using these electronic clinical information retrieval tools. Nevertheless, there was a lot of variation in the level of knowledge before initiating the search, and all residents much improved the proportion of their correct answers after the search (data not shown), indicating that most clinicians should be able to improve their performance at finding answers to clinical questions using either search engine. On the other hand, using a sample of family medicine residents was not a

threat to internal validity considering the experimental randomized design of our study.

The fourth limitation concerns the study's experimental setting. Participants did not answer questions at the point of care. They were also allowed to search longer than the 2 minutes estimated to be feasible during clinical practice to answer 1 question.²⁹ We accepted this limitation in light of our aim to test the search engines in a “reflection on action” mode (after the physician-patient encounter) rather than in a “reflection in action” mode (during the encounter). Having now demonstrated that InfoClinique and the Trip database can be used to find correct answers to a large proportion of clinical questions typical in family medicine, we believe that it would be appropriate to evaluate the efficacy of these tools during patient encounters.³⁰⁻³²

Fifth, some results of the initial searches were not saved owing to problems with the study website. This was not associated with the ability of residents to save their searches or to the fact that an answer was correct or not. These represented only 5% (15 of 300) of all possible results of the first searches. We do not believe this substantially biased our results.

Sixth, this study was conducted between February and May 2007. The 2 search engines have since been modified to reflect the evolution of available Web-based resources, and it is possible that the resources they currently index would produce different results than those produced in 2007. However, to our knowledge neither engine has undergone a substantial change that would have decreased users' ability to find correct answers. If anything, their current interface and the way that they now index websites should have improved searching capacity and the effect on clinical decision making.

Our results and limitations indicate that more evaluations and comparisons of clinical information retrieval tools are needed. Clinical questions should cover a larger span of clinical themes that discuss diagnosis, pathogenesis, and prognosis in addition to therapy. Studies should be conducted both during and after patient encounters. They should rely on objective outcome measures and include baseline data. Studies on the validation of instruments such as the CI-SDM, UCIS, and SORT should also take place.

We also acknowledge that finding the best available medical evidence with a medical search engine does not automatically improve practice and patient care. Based on a Cochrane systematic review, there is currently insufficient evidence to support or refute the use of any electronic clinical information retrieval tools by health care providers to this end.³³ However, finding such information is essential if evidence-based practice and shared decision making are to occur in clinical practice. Physicians have difficulty estimating the benefits and harms of interventions they commonly

prescribe.³⁴ Using effective medical search engines to access pre-appraised evidence should help clinicians apply this evidence in practice. Ideally, the engines would also index patient decision aids aimed at fostering shared decision making.^{34,35}

Conclusion

This trial showed that both InfoClinique and the Trip database provided access to evidence-based clinical information on the benefits and harms of treatment and preventive interventions used in family medicine, information which is a prerequisite to evidence-based practice and shared decision making. The choice of one search engine over the other thus becomes a matter of preference.

At the time of the study, **Dr Labrecque** worked in the Research Centre of the Centre hospitalier universitaire de Québec (CHUQ) in Québec city, and in the Department of Family and Emergency Medicine at Laval University in Québec city. **Mr Ratté** worked in the Research Centre of the CHUQ. **Dr Frémont** worked in the Research Centre of the CHUQ and in the Department of Rehabilitation at Laval University. **Drs Cauchon** and **Ouellet** worked in the Department of Family and Emergency Medicine at Laval University. **Drs Hogg** and **McGowan** worked in the Department of Family Medicine at the University of Ottawa in Ontario. **Dr Gagnon** worked in the Research Centre of the CHUQ and in the Faculty of Nursing at Laval University. **Mr Njoya** worked in the Research Centre of the CHUQ. **Dr Légaré** worked in the Research Centre of the CHUQ and in the Department of Family and Emergency Medicine at Laval University.

Acknowledgment

This study was funded by the Fonds Gilles-Cormier from Laval University and the College of Family Physicians of Canada (Janus grant). **Dr Légaré** is Tier 2 Canada Research Chair in Implementation of Shared Decision Making in Primary Care. We thank **Mehdi Atmani** for the development of the Web-based platform and **Doug Salwedel** who provided answers following the just-in-time protocol. We thank **Jennifer Petrela** for editing the paper.

Contributors

Drs Labrecque, Légaré, Cauchon, Frémont, Hogg, McGowan, and Gagnon wrote the research protocol. All authors approved the final protocol. **Drs Labrecque, Frémont, and Cauchon** and **Mr Ratté** were responsible for the overall conduct of the study. **Mr Ratté** was the project's research coordinator. **Drs Labrecque, Cauchon, Frémont, Ouellet, and Hogg** and **Mr Ratté** participated in data collection. **Drs Labrecque** and **Légaré** and **Mr Ratté** participated in data analyses and interpretation. **Mr Ratté** and **Dr Labrecque** drafted the first version of the manuscript. All authors approved the final manuscript.

Competing interests

Drs Labrecque, Cauchon, and Frémont are members of the development team of InfoClinique. Participants were not offered any financial incentives. Use of InfoClinique and the Trip database is free.

Correspondence

Dr Michel Labrecque, Laval University, Département de médecine familiale et de médecine d'urgence, Hôpital St-François d'Assise, D6-728, 10 rue de l'Espinay, Québec city, QC G1L 3L5; telephone 418 525-4444, extension 52419; fax 418 525-4194; e-mail michel.labrecque@mfa.ulaval.ca

References

1. Straus SE, Richardson WS, Glasziou P, Haynes B. *Evidence based medicine: how to practice and teach EBM*. 3rd ed. New York, NY: Elsevier; 2005.
2. Makoul G, Clayman ML. An integrative model of shared decision making in medical encounters. *Patient Educ Couns* 2006;60(3):301-12.
3. Montori VM, Guyatt GH. Progress in evidence-based medicine. *JAMA* 2008;300(15):1814-6.
4. Elwyn G, Laitner S, Coulter A, Walker E, Watson P, Thomson R. Implementing shared decision making in the NHS. *BMJ* 2010;341:c5146.
5. Dicenso A, Bayley L, Haynes RB. Accessing pre-appraised evidence: fine-tuning the 5S model into a 6S model. *Evid Based Nurs* 2009;12(4):99-101.
6. Banzi R, Cinquini M, Liberati A, Mocschetto I, Pecoraro V, Taqilabue L, et al. Speed of updating online evidence based point of care summaries: prospective cohort analysis. *BMJ* 2011;343:d5856.
7. Trip [database online]. Trip Database Ltd; 2013. Available from: www.tripdatabase.com. Accessed 2010 Dec 5.

8. Coumou HC, Meijman FJ. How do primary care physicians seek answers to clinical questions? A literature review. *J Med Libr Assoc* 2006;94(1):55-60.
9. Pluye P, Grad RM, Dunikowski LG, 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.
10. 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.
11. Patel MR, Schardt CM, Sanders LL, Keitz SA. Randomized trial for answers to clinical questions: evaluating a pre-appraised versus a MEDLINE search protocol. *J Med Libr Assoc* 2006;94(4):382-7.
12. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993;270(21):2598-601.
13. InfoClinique [database online]. Québec city, QC: Laval University; 2013. Available from: <http://infoclinique.fmed.ulaval.ca>. Accessed 2010 Dec 5.
14. Brassey J, Elwyn G, Price C, Kinnersley P. Just in time information for clinicians: a questionnaire evaluation of the ATTRACT project. *BMJ* 2001;322(7285):529-30.
15. Meats E, Brassey J, Heneghan C, Glasziou P. Using the Turning Research Into Practice (Trip) database: how do clinicians really search? *J Med Libr Assoc* 2007;95(2):156-63.
16. McGowan J, Hogg W, Campbell C, Rowan M. Just-in-time information improved decision-making in primary care: a randomized controlled trial. *PLoS ONE* 2008;3(11):e3785.
17. Legare F, Graham I, O'Connor A, Dolan JG, Belanger-Ducharme F. Prise de décision partagée: traduction et validation d'une échelle de confort décisionnel du médecin. *Pédagogie Médicale* 2003;4(4):216-22.
18. Shaughnessy AF, Slawson DC, Bennett JH. Becoming an information master: a guidebook to the medical information jungle. *J Fam Pract* 1994;39(5):489-99.
19. Grad RM, Pluye P, Meng Y, Segal B, Tamblin R. Assessing the impact of clinical information-retrieval technology in a family practice residency. *J Eval Clin Pract* 2005;11(6):576-86.
20. Pluye P, Grad RM, Stephenson R, Dunikowski LG. A new impact assessment method to evaluate knowledge resources. *AMIA Annu Symp Proc* 2005:609-13.
21. Ebell MH, Siwek J, Weiss BD, Woolf SH, Susman J, Ewigman B, et al. Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature. *J Am Board Fam Pract* 2004;17(1):59-67.
22. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process* 1991;50(2):179-211.
23. Venkatesh V, Davis F. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manage Sci* 2000;46(2):186-204.
24. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
25. Hersh WR, Crabtree MK, Hickam DH, Sacherek L, Rose L, Friedman CP. Factors associated with successful answering of clinical questions using an information retrieval system. *Bull Med Libr Assoc* 2000;88(4):323-31.
26. Hersh WR, Crabtree MK, Hickam DH, Sacherek L, Friedman CP, Tidmarsh P, et al. Factors associated with success in searching MEDLINE and applying evidence to answer clinical questions. *J Am Med Inform Assoc* 2002;9(3):283-93.
27. Westbrook JI, Coiera EW, Gosling AS. Do online information retrieval systems help experienced clinicians answer clinical questions? *J Am Med Inform Assoc* 2005;12(3):315-21.
28. Gorman PN, Helfand M. Information seeking in primary care: how physicians choose which clinical questions to pursue and which to leave unanswered. *Med Decis Making* 1995;15(2):113-9.
29. Ely JW, Osheroff JA, Ebell MH, Bergus GR, Levy BT, Chambliss ML, et al. Analysis of questions asked by family doctors regarding patient care. *BMJ* 1999;319(7206):358-61.
30. Schön DA. *The reflective practitioner: how professionals think in action*. New York, NY: Basic Books; 1983.
31. Schön DA. *Educating the reflective practitioner*. San Francisco, CA: Jossey Bass; 1987.
32. Lockyer J, Gondocz ST, Thivierge RL. Knowledge translation: the role and place of practice reflection. *J Contin Educ Health Prof* 2004;24(1):50-6.
33. McGowan JL, Grad R, Pluye P, Hannes K, Deane K, Labrecque M, et al. Electronic retrieval of health information by healthcare providers to improve practice and patient care. *Cochrane Database Syst Rev* 2009;(3):CD004749.
34. Desbiens G, Labrecque M, Bergeron ME, Larivière ME, Carlos Ochoa J, Njoya M. Do family physicians correctly estimate the benefits and risks of common therapeutic and preventive intervention? Poster presented at Family Medicine Forum; 2010 Oct 14-16; Vancouver, BC.
35. Elwyn G, O'Connor A, Stacey D, Volk R, Edwards A, Coulter A, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ* 2006;333(7565):417.
