



Published in final edited form as:

Int J Med Inform. 2016 February ; 86: 126–134. doi:10.1016/j.ijmedinf.2015.11.006.

Formative evaluation of a patient-specific clinical knowledge summarization tool

Guilherme Del Fiol^{a,*}, Javed Mostafa^b, Dongqiuye Pu^b, Richard Medlin^b, Stacey Slager^a, Siddhartha R. Jonnalagadda^c, and Charlene R. Weir^{a,d}

^aDepartment of Biomedical Informatics, University of Utah, Salt Lake City, UT, USA

^bSchool of Information and Library Science (SILS), University of North Carolina, Chapel Hill, NC, USA

^cDivision of Health and Biomedical Informatics, Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

^dDepartment of Veterans Affairs Medical Center, Salt Lake City, UT, USA

Abstract

Objective—To iteratively design a prototype of a computerized clinical knowledge summarization (CKS) tool aimed at helping clinicians finding answers to their clinical questions; and to conduct a formative assessment of the usability, usefulness, efficiency, and impact of the CKS prototype on physicians' *perceived decision quality* compared with standard search of UpToDate and PubMed.

Materials and methods—Mixed-methods observations of the interactions of 10 physicians with the CKS prototype vs. standard search in an effort to solve clinical problems posed as case vignettes.

Results—The CKS tool automatically summarizes patient-specific and actionable clinical recommendations from PubMed (high quality randomized controlled trials and systematic reviews) and UpToDate. Two thirds of the study participants completed 15 out of 17 usability tasks. The median time to task completion was less than 10 s for 12 of the 17 tasks. The difference in search time between the CKS and standard search was not significant (median = 4.9 vs. 4.5 min). Physician's *perceived decision quality* was significantly higher with the CKS than with manual search (mean = 16.6 vs. 14.4; $p = 0.036$).

*Corresponding author at: Department of Biomedical Informatics, University of Utah, 421 Wakara Way, Salt Lake City, UT 84108, USA. Fax: +1 801 581 4297. guilherme.delfiol@utah.edu (G. Del Fiol).

Conflicts of interest

The authors have no conflicts of interest to declare.

Authors' contributions

GDF, JM, DP, RM, SRJ, and CRW contributed to the conception and design of the study. GDF, JM, DP, RM, SS, and CRW contributed to the acquisition of data. GDF, JM, SS, and CRW contributed to the analysis and interpretation of data. GDF drafted the article. All co-authors revised the article critically for important intellectual content and provided final approval of the version to be submitted.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ijmedinf.2015.11.006>.

Conclusions—The CKS prototype was well-accepted by physicians both in terms of usability and usefulness. Physicians perceived better decision quality with the CKS prototype compared to standard search of PubMed and UpToDate within a similar search time. Due to the formative nature of this study and a small sample size, conclusions regarding efficiency and efficacy are exploratory.

Keywords

Clinical decision support; Information seeking and retrieval; Online information resources; Information needs

1. Introduction

Clinicians often raise clinical questions in the course of patient care and are unable to find answers to a large percentage of these questions. A systematic review of 21 studies has shown that clinicians raised on average one clinical question out of every two patients seen and that over half of these questions were left unanswered [1]. Recent advances in online clinical knowledge resources offer an opportunity to address this problem. Studies have shown that, when used, these resources are able to answer over 90% of clinicians' questions, improving clinicians' performance and patient outcomes [2–12].

Despite increased clinician adoption of online resources, recent studies still show a stable picture, with most questions continuing to be left unanswered [1]. It appears that significant barriers still challenge the use of online resources to support clinical decisions [1]. One possibility is that clinicians often cannot process the large amount of potentially relevant information available within the timeframe of a typical patient care setting. Several solutions have been designed to address this problem, such as manual curation and synthesis of the primary literature (e.g., systematic reviews, clinical guidelines, evidence summaries), context-specific links to relevant evidence resources within electronic health record (EHR) systems [10], and clinical question answering systems [13]. Although some of these solutions have shown to be effective, there are still opportunities for further improvement [10,14,15].

In the present research we explore automatic text summarization and information visualization techniques to design a clinical decision support (CDS) tool called Clinical Knowledge Summary (CKS). Given the tight timelines and competing demands on the attention of providers, designing a system for knowledge summarization is extremely important. Currently, allotted time to see patients has been decreasing with more intense pressure to expand revenues. As a result, primary care clinicians in outpatient care may only have less than a few minutes to pursue questions. The CKS retrieves and summarizes patient-specific, actionable recommendations from PubMed citations and UpToDate articles. The summarization output is presented to users in a highly interactive display. In the present study we report the results of a formative, mixed-methods assessment of a high-fidelity prototype of the CKS. The study aimed to assess the usability of the CKS, obtain insights to guide CKS design, and assess the CKS impact on physicians ability to solve questions in case vignettes.

2. Methods

The study consisted of mixed-methods observations of physician interactions with the CKS prototype in an effort to solve clinical problems posed as case vignettes. The study addressed the following research questions: (1) To what degree are the CKS features *easy to use* and *useful*; (2) how *efficient* are CKS searches as compared to manual searches; and (3) how do CKS searches differ from manual searches in terms of clinician's *perceived decision quality*?

2.1. CKS tool design

The CKS design was guided by the following set of principles derived from Information Foraging theory [16]: (i) maximize *information scent* (i.e., cues to help identify relevant information); (ii) facilitate the cost-benefit assessment of *information-seeking effort* by providing measures of the amount of information available and enabling quick information scanning; and (iii) enable *information patch* enrichment (i.e., features that allow users to increase the concentration of relevant content). We also followed Shneiderman's visual information seeking principles: (i) first, information overview from each source; (ii) followed by zoom and filtering; and (iii) then on-demand access to details [17].

The CKS was designed in rapid iterative cycles guided by feedback and insights obtained from informal user interactions with prototypes. In the early cycles, we experimented with multiple alternate "low-fidelity" prototypes in the form of diagrams and screen mockups. The low-fidelity prototypes progressively evolved towards "high-fidelity" prototypes implemented in HTML and JavaScript until a more stable design was achieved for the formative evaluation. During the formative evaluation, the CKS tool went through one additional version to improve usability after exposure to the first set of study participants. Fig. 1 depicts the method employed in each CKS design stage.

2.2. CKS software architecture

The CKS architecture consists of two independent processes (Fig. 2) built over open source and publicly available components. To enable real-time performance for the CKS, we pre-process text sources through a text summarization pipeline and store the results in a relational database (Fig. 2A). High quality clinical studies are identified from PubMed using a machine learning classifier developed by Kilicoglu et al. (Fig. 2A, Step A.1) [18]. PubMed abstracts and UpToDate articles are processed by a classifier that uses concepts, semantic predications, and deontic terms as predictors of sentences that provide clinically actionable recommendations (Step A.2) [19,20]. The output from Steps A.1 and A.2 is stored in a relational database in the form of sentence-level metadata (Step A.3).

At real-time, the CKS application, which was developed in HTML and JavaScript, starts the process by sending a Hypertext Transfer Protocol (HTTP) request to OpenInfobutton [21], a Java-based, Infobutton Manager Web service compliant with the Health Level Seven (HL7) Infobutton Standard (Fig. 2B, Step B.1) [22]. The request includes contextual information about the patient, the user, and the care setting.

The Infobutton Manager uses this information to retrieve relevant articles from PubMed and UpToDate using PubMed's native search engine and UpToDate's HL7 Infobutton Web service (Step B.2). Relevant documents are retrieved in JSON format, also compliant with the HL7 Infobutton Standard (Step B.3). For the retrieved articles, the Infobutton Manager uses standard query language (SQL) to retrieve a subset of high quality articles along with clinically actionable sentences from the pre-processed summarization database (Step B.4). These recommendations along with links to the original source are aggregated to produce an output in JSON format compliant with the HL7 Infobutton Standard (Step B.5). The CKS user interface parses the JSON output and presents the clinically actionable statements extracted from UpToDate articles and PubMed abstracts. Features of the CKS user interface are described in Section 3. Details of the CKS architecture, the HL7 Infobutton Standard, and the summarization algorithms are described elsewhere [18–21,23–25].

2.3. Study settings

Formative evaluation sessions were conducted at the University of Utah and the University of North Carolina Chapel Hill. Sessions were conducted onsite and remotely via online conference meeting. Participants accessed an instance of the CKS that was hosted in the cloud.

2.4. Participants

Users in the iterative design stage included clinician collaborators and members of the research team. For the formative evaluation, we recruited a sample of 10 physicians who had not participated in the iterative design stage and had no previous exposure to the tool. We sought a purposive sample of physicians with various specialties and a wide range of clinical experience. The goal was to expose the tool to a diverse group of users.

2.5. Case vignettes

We adapted six case vignettes that were developed and validated in previous studies [26–31]. All vignettes were focused on patient treatment and covered a range of medical problems in different areas, such as diabetes mellitus, atrial fibrillation, and depression. To increase the complexity of the cases, the problem posed in the vignettes could be resolved in multiple ways. The goal was to stimulate physicians to consider multiple treatment alternatives in their information-seeking sessions.

2.6. Procedure

Sessions started with a brief introduction and description of the study. Participants interacted with a brief, 2-slide tutorial describing the CKS tool. Next, participants were assigned to three case vignettes, each of which was used in a specific segment of the study session.

2.6.1. CKS usability—The first segment was focused on CKS usability and allowed participants to familiarize themselves with the tool. Participants were asked to complete 17 tasks within the CKS (Table 1), such as finding a relevant randomized controlled trial or systematic review, finding a study sample size and funding source, and linking to the original source of a particular sentence. The tasks were designed to cover all the

functionality available in the CKS tool. Participants were asked to “think-aloud” as they worked on these tasks to identify usability issues and to obtain feedback for improvement of the CKS tool.

2.6.2. Clinical problem-solving—The second and third segments were focused on solving a clinical problem described in the vignettes. Participants were asked to use the CKS in one of the problem-solving segments and UpToDate/PubMed directly (manual search) in the other segment in random order. Therefore, every participant was exposed to the usability segment, manual search, and CKS. For the CKS sessions, clinical information from the vignettes was transmitted to the CKS Web service via HL7 Infobutton Standard. After the manual search and CKS segments, participants were asked to complete a post-session questionnaire (available as online supplement). The questionnaire assessed: (i) self-perceived impact of the information found on providers’ decisions, knowledge, recall, uncertainty, confidence, information-seeking effort and likelihood to refer the patient to a specialist according to a Likert scale; and (ii) perceived usefulness of each of the CKS features. The types of metrics used to measure impact in the questionnaire were based on previous research [32].

On-site sessions were recorded with screen capture software (Hypercam). Remote sessions were recorded with Web meeting software (Webex).

2.7. Data analysis

Video recordings were imported into qualitative analysis software (*ATLAS.ti*) for coding purposes. The usability segments were coded according to fragments in which each usability task was performed. Subsequently, each usability task was coded according to (i) successful completion; (ii) time to completion; (iii) implicit non-verbalized usability issues (e.g., pauses, hesitation, mouse movement indicating user was lost); and (iv) verbalized issues identified in participants’ think aloud statements (e.g., requests for help, provision of feedback). Content analysis was performed on the implicit and explicit (think aloud) user responses to identify usability issues and areas for improvement.

The manual search and CKS segments were coded according to (i) session duration; (ii) content accessed (randomized controlled trials, systematic reviews, UpToDate); and (iii) use of specific CKS features (e.g., filters, link to original source).

The quantitative analysis was largely descriptive and exploratory. For use of CKS features, we calculated the average, standard deviation, and median for the time spent on each feature and the percentage of physicians who used each feature. To test for a differential impact of vignette type, a within-subject *t*-test was conducted for vignette perceived complexity and personal experience in the vignette’s domain. To test for differential impact of type of search, we conducted a within-subject analysis (i.e., paired *t*-test for repeated measures) aggregating across vignettes on *information-seeking time* and the main cognitive variable of *perceived decision quality*. The latter was composed of the sum of physician Likert scale ratings of *increase in knowledge*, *enhanced decision-making*, *increased recall*, and *enhanced confidence with the decision-making process*. Adjustments for multiple comparisons were not made because there was only one comparison per hypothesis and those comparisons

were planned a priori. To explore the possible affective mechanisms associated with *perceived decision quality*, partial correlations (controlling for vignette experience) were assessed between the *perceived decision quality* variable and *improved recall*, *surprise*, *effort*, *frustration* and *likelihood to refer*.

3. Results

The results section first presents findings related to the design of the CKS user experience, followed by usability testing results, and an exploration of the impact of the two search methods on the pre-identified cognitive factors.

3.1. CKS user experience

The design of the CKS user experience required eight design iterations that were carried out before the formative evaluation. Fig. 3 shows a screen shot of the CKS tool at the end of the iterative design stage. Fig. 4 shows the fine-tuned version after exposure to the first subset of study participants in the formative evaluation. The resulting CKS tool generates interactive, patient-specific summaries of the recent clinical evidence available in UpToDate and PubMed, organized into three boxes: (i) summarized articles from UpToDate; (ii) excerpts of recent high impact systematic reviews (SR) from PubMed; (iii) and excerpts of recent high impact randomized controlled trials (RCT) from PubMed. Filters allow the user to focus a summary on a specific treatment of interest. Summary sentences can be expanded to display surrounding sentences from the original source. Links allow navigating from specific sentences to their original source within an UpToDate article or a PubMed abstract.

3.2. CKS improvements during the formative evaluation

Observations of the initial physician interactions with the CKS in the formative evaluation showed that the main opportunity for design improvement was in the CKS landing page. In the original version, the landing page displayed summaries of a set of retrieved UpToDate articles. However, we observed that physicians would often focus on only one article and that the article title was the key to determining its relevancy and subsequent information-seeking strategy. Therefore, rather than displaying summaries for all retrieved UpToDate articles (Fig. 3), the revised version displays only article titles along with the number of sentences retrieved (Fig. 4). A specific article summary can be accessed by clicking on the article title. This approach presents CKS users with the minimum information necessary (i.e., article titles and number of relevant sentences available) to help define their initial information seeking strategy. This initial assessment is equivalent to the cost-benefit assessment of information patches as described in the Information Foraging theory.

Other specific areas in the user experience were also fine-tuned. The ability to expand a summary sentence to display adjoining sentences was rated as a useful feature. In the original version, this feature was triggered by hovering the mouse cursor over a sentence. Study participants complained that this mechanism was oversensitive, causing adjacent sentences to be unintentionally displayed. In the redesigned version, this feature was redesigned as a hyperlink, marked as “more,” which was positioned at the end of each sentence (Fig. 4). This approach is a common design feature adopted by numerous Web

sites, such as those that provide user reviews. Filters were also redesigned based on common Web features, especially those used in e-commerce Web sites.

Also in the original version, users could click on an UpToDate sentence to navigate to the full article in UpToDate. However, we observed that some physicians highlighted sentences they read and this would unintentionally trigger a link to UpToDate. To address this issue, we relocated this feature to an UpToDate logo that was placed at the end of each section heading (Fig. 4).

3.3. Ease of use and usefulness of CKS features

3.3.1. Timing and completion—Two thirds of the participants completed 15 out of the 17 usability tasks (Table 1 and Fig. 5). The median time to completion was less than 10 s for 12 of the 17 tasks. Two tasks were not completed by over 50% of the participants and took a median of 16 s to complete: finding the sample size and source of funding of an RCT.

In the CKS sessions, all physicians used the UpToDate excerpts box, spending a median of 3.3 min on this feature (Table 2 and Fig. 6). From a sentence in the UpToDate box, 70% of the physicians navigated to the original source in UpToDate for further details. While 60% of the participants accessed RCTs and systematic reviews when using the CKS, only 20% did so when searching PubMed and UpToDate manually (Fig. 6)

3.3.2. Ratings of feature usefulness of the CKS—The links from specific sentences to their source within an article in the UpToDate Web site was the highest rated feature of the CKS, followed by treatment filters and the ability to display sentences adjacent to a summary sentence (Table 3).

3.4. CKS vs. manual search comparisons

3.4.1. Vignette differences—Ratings of perceived vignette complexity and the individual's prior experience in the vignette domain were not statistically significant using a within-subject repeated measures *t*-test. These results suggest that vignette complexity and difficulty played a minimal role.

3.4.2. Timing for searches—Physicians spent a median of 4.9 min (range 1.1–11.8) seeking information with the CKS vs. 4.5 with manual search (range 2.0–10.4). This difference was not statistically significant.

3.4.3. Decision-making variables—A composite variable was created that combined the four *perceived decision quality* variables in order to maximize reliability, generalizability and to minimize excessive comparisons. This variable was the sum of ratings of perceived *increase in knowledge*, *enhanced decision-making*, *increased recall*, and *enhanced confidence with the decision-making process*. The Cronbach's alpha for the two scales (manual search and CKS search responses) were 0.81 and 0.64, respectively. In a comparison of physician's *perceived decision quality*, ratings were significantly higher with the CKS than with manual search (mean = 16.6 vs. 14.4; $t_9 = 2.47$; $p = 0.036$). The effect size obtained was 0.78, which by definition according to Cohen, is a moderate effect size [33]. No clear-cut possible mediators emerged in exploration of the correlations within

groups, except *perceived frustration*, which was marginally correlated with *perceived decision quality* in both interventions ($r = 0.60, p = 0.07$ and $0.63, p = 0.09$ for manual search and CKS searches respectively). Although a mediation analysis is beyond the scope of this exploratory study, *perceived frustration* might be one mechanism (Table 4).

4. Discussion

Despite advances in information retrieval technology and demonstrated impact of online clinical evidence resources on patient care [14], barriers still limit the use of these resources to support clinical decision-making [1]. The overall goal of our research is to lower some of these barriers by reducing cognitive effort in evidence seeking for complex treatment questions. Previous approaches to lowering barriers to clinicians' use of online evidence for CDS include automating the selection of evidence resources and the search process based on the clinical context within the EHR [10]; searching multiple resources simultaneously through federated search tools [11,34]; and clinical question answering systems [13]. Our approach borrows some of the methods in these previous approaches, such as using the clinical context to formulate automated searches in multiple evidence sources; and extracting most useful and relevant sentences from these resources. However, previous summarization systems for CDS produce a static summary [35]. Unlike these previous approaches, the CKS produces dynamic and interactive summaries that can be narrowed to specific treatment topics based on the users' information needs. From a methods perspective, our overall approach consisted of adapting and integrating recent advances in biomedical text summarization, guided by information foraging theory, information visualization principles, and empirical user-centered design. This multi-faceted approach is needed for integrating advances in biomedical text summarization into CDS tools that can be used in real patient care settings. The results of our early, formative evaluation are promising and suggest increased *perceived decision quality* (16.6 vs. 14.4 out of 20) without reduction in efficiency.

According to the technology acceptance model (TAM), ease of use and perceived usefulness of a technology are predictors of actual use [36]. In our study, participants completed most of the usability tasks in a relatively short amount of time and with minimal training. Moreover, most CKS features were highly rated and often used by physicians in the clinical problem-solving sessions. Several factors may have contributed to these findings, including the user-centered, iterative design process, informed by incremental low and high-fidelity prototypes, and guided by theoretical principles and design strategies.

The clinical-problem solving sessions suggest that the CKS promoted larger consumption of evidence from the primary literature when compared to manual search. This is an important finding, since clinicians often perceive primary literature resources, such as PubMed, to be too difficult to use in busy patient care settings [37]. Design features focused on specific concepts in information foraging theory may have contributed to reducing barriers to accessing evidence in the primary literature: (1) automated search for recent, high impact studies and context-specific filters to enable *information patch enrichment*, i.e., ready access to an increased rate of relevant and high-quality information; (2) extraction of key study information to increase *information scent*; and (3) incremental access to details to encourage

information patch exploration. Collectively, these design features may have contributed to reducing the perceived cost and increasing the perceived value of pursuing information from multiple information patches. Further improvements could be obtained in the extraction of key pieces of information, such as those included in the PICO format (population, intervention, comparison, and outcome) [38].

Most participants (70%) navigated from UpToDate excerpts in the CKS to the original source in UpToDate. This finding supports one of our design goals for the CKS to be a complimentary tool, rather than a replacement of existing tools, that enables access to summarized evidence from multiple sources.

Unlike most common approaches, our study went beyond usability testing to explore possible impact of the CKS on both cognitive and affective measures. This comprehensive approach sets the stage for a deeper understanding of the impact of design in the context of real-world settings. The formative evaluation supports the efficacy of the design at several levels. Physicians using the CKS vs. manual search perceived better decision quality, with no impact on efficiency. Despite the early design stage of the CKS, it is possible that physicians using the tool were able to scan through a larger number of alternatives, as compared to a full narrative document, without compromising efficiency.

4.1. Limitations

The main limitation of this study is the relatively small sample size for the quantitative assessments. Thus, quantitative comparisons are exploratory and study conclusions are not definite. Nevertheless, the sample size is considered adequate for usability assessments [39] and the quantitative measures provide exploratory comparisons and effect size estimate for future similar studies. To reach a power of 0.8 for this effect size, future studies should aim for a sample size of 16 subjects. Limitations of the quantitative inferences also include the likely interference of the think-aloud method with physicians' cognition and a research environment that promoted exploration rather than efficiency. Physicians did not have previous experience with the CKS tool and were encouraged to explore the available features. Curiosity regarding the new tool and physicians' lack of familiarity with it may have contributed to longer CKS sessions.

5. Conclusion

We describe the design and formative evaluation of a clinical decision support tool that automatically summarizes patient-specific evidence to help clinicians' decision making. We followed an iterative, user-centered design, guided by information foraging theory and information visualization principles. High-fidelity prototypes allowed target users to interact with incremental versions of the tool and provide feedback. Results of the formative study showed that overall physicians were able to complete most of the usability tasks in a short period of time. The feature most often used by study participants was the clinical recommendations extracted from UpToDate. Unlike manual search sessions, physicians with the CKS often reviewed information from randomized controlled trials and systematic reviews. Physicians' *perceived decision quality* was significantly higher with the CKS than with manual search, but no difference was found in information seeking time. Due to the

formative nature of this study and a small sample size, conclusions regarding efficiency and efficacy are exploratory. A summative study is being planned with a larger sample size and procedures that aim to simulate a patient care environment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This project was supported by grants 1R01LM011416-02 and 4R00LM011389-02 from the National Library of Medicine. The funding source had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

References

1. Del Fiol G, Workman TE, Gorman PN. Clinical questions raised by clinicians at the point of care: a systematic review. *JAMA Intern. Med.* 2014; 174:710–718. [PubMed: 24663331]
2. Chambliss ML, Conley J. Answering clinical questions. *J. Fam. Pract.* 1996; 43:140–144. [PubMed: 8708623]
3. Crowley SD, Owens TA, Schardt CM, Wardell SI, Peterson J, Garrison S, Keitz SA. A Web-based compendium of clinical questions and medical evidence to educate internal medicine residents. *Acad. Med.* 2003; 78:270–274. [PubMed: 12634206]
4. Schilling LM, Steiner JF, Lundahl K, Anderson RJ. Residents' patient-specific clinical questions: opportunities for evidence-based learning. *Acad. Med.* 2005; 80:51–56. [PubMed: 15618093]
5. Green ML, Ciampi MA, Ellis PJ. Residents' medical information needs in clinic: are they being met? *Am. J. Med.* 2000; 109:218–223. [PubMed: 10974185]
6. Del Mar CB, Silagy CA, Glasziou PP, Weller D, Spinks AB, Bernath V, Anderson JN, Hilton DJ, Sanders SL. Feasibility of an evidence-based literature search service for general practitioners. *Med. J. Aust.* 2001; 175:134–137. [PubMed: 11548078]
7. Ette EI, Achumba JI, Brown-Awala EA. Determination of the drug information needs of the medical staff of a Nigerian hospital following implementation of clinical pharmacy services. *Drug Intell. Clin. Pharm.* 1987; 21:748–751. [PubMed: 3652937]
8. Swinglehurst DA, Pierce M, Fuller JC. A clinical informaticist to support primary care decision making. *Qual. Health Care.* 2001; 10:245–249. [PubMed: 11743154]
9. Verhoeven AA, Schuling J. Effect of an evidence-based answering service on GPs and their patients: a pilot study. *Health Inf. Lib. J.* 2004; 21(Suppl. 2):27–35.
10. Del Fiol G, Haug PJ, Cimino JJ, Narus SP, Norlin C, Mitchell JA. Effectiveness of topic-specific infobuttons: a randomized controlled trial. *J. Am. Med. Inf. Assoc.* 2008; 15:752–759.
11. Magrabi F, Coiera EW, Westbrook JI, Gosling AS, Vickland V. General practitioners' use of online evidence during consultations. *Int. J. Med. Inf.* 2005; 74:1–12.
12. Sackett DL, Straus SE. Finding and applying evidence during clinical rounds: the evidence cart. *JAMA.* 1998; 280:1336–1338. [PubMed: 9794314]
13. Cao Y, Liu F, Simpson P, Antieau L, Bennett A, Cimino JJ, Ely J, Yu H. AskHERMES: an online question answering system for complex clinical questions. *J. Biomed. Inform.* 2011; 44:277–288. [PubMed: 21256977]
14. 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. Inf.* 2005; 74:745–768.
15. Maviglia SM, Yoon CS, Bates DW, Kuperman G. KnowledgeLink: impact of context-sensitive information retrieval on clinicians' information needs. *J. Am. Med. Inf. Assoc.* 2006; 13:67–73.
16. Pirolli P, Card S. Information foraging. *Psychol. Rev.* 1999; 106:643–675.

17. Shneiderman, B. The eyes have it: a task by data type taxonomy for information visualizations, visual languages, 1996; Proceedings., IEEE Symposium on, IEEE; 1996. p. 336-343.
18. Kilicoglu H, Demner-Fushman D, Rindflesch TC, Wilczynski NL, Haynes RB. Towards automatic recognition of scientifically rigorous clinical research evidence. *J. Am. Med. Inf. Assoc.* 2009; 16:25–31.
20. Morid MA, Jonnalagadda S, Fiszman M, Raja K, Del Fiol G. Classification of Clinically Useful Sentences in MEDLINE. American Medical Informatics Association (AMIA) Annual Symposium. 2015; 2015
21. Del Fiol G, Curtis C, Cimino JJ, Iskander A, Kalluri AS, Jing X, Hulse NC, Long J, Overby CL, Schardt C, Douglas DM. Disseminating context-specific access to online knowledge resources within electronic health record systems. *Stud. Health Technol. Inf.* 2013; 192:672–676.
22. Del Fiol G. Context-aware Knowledge Retrieval (Infobutton), Knowledge Request Standard. Health Level 7 International. 2010
23. Mishra R, Del Fiol G, Kilicoglu H, Jonnalagadda S, Fiszman M. Automatically extracting clinically useful sentences from UpToDate to support clinicians' information needs. *AMIA Annu. Symp. Proc.* 2013; 2013:987–992. [PubMed: 24551389]
24. Zhang M, Del Fiol G, Grout RW, Jonnalagadda S, Medlin R Jr, Mishra R, Weir C, Liu H, Mostafa J, Fiszman M. Automatic identification of comparative effectiveness research from medline citations to support clinicians' treatment information needs. *Stud. Health Technol. Inf.* 2013; 192:846–850.
25. Jonnalagadda SR, Del Fiol G, Medlin R, Weir C, Fiszman M, Mostafa J, Liu H. Automatically extracting sentences from Medline citations to support clinicians' information needs. *J. Am. Med. Inf. Assoc.* 2013; 20:995–1000.
26. Swennen MH, Rutten FH, Kalkman CJ, van der Graaf Y, Sachs AP, van der Heijden GJ. Do general practitioners follow treatment recommendations from guidelines in their decisions on heart failure management? A cross-sectional study. *BMJ Open.* 2013; 3:e002982.
27. Resnick B. Atrial fibrillation in the older adult. Presentation and management issues. *Geriatric Nurs. (New York, N.Y.).* 1999; 20:188–193. quiz 194.
28. Falchi A, Lasserre A, Gally A, Blanchon T, Sednaoui P, Lassau F, Massari V, Turbelin C, Hanslik T. A survey of primary care physician practices in antibiotic prescribing for the treatment of uncomplicated male gonococcal urethritis. *BMC Fam. Pract.* 2011; 12:35. [PubMed: 21592343]
29. Drexel C, Jacobson A, Hanania NA, Whitfield B, Katz J, Sullivan T. Measuring the impact of a live, case-based, multiformat, interactive continuing medical education program on improving clinician knowledge and competency in evidence-based COPD care. *Int. J. Chron. Obstruct. Pulmon. Dis.* 2011; 6:297–307. [PubMed: 21697994]
30. Hirsch IB, Molitch ME. Glycemic Management in a Patient with Type 2 Diabetes. *New England J. Med.* 2013; 369:1370–1372. [PubMed: 24088099]
31. Treatment of a 6-Year-Old Girl with Vesicoureteral Reflux. *New England Journal of Medicine.* 2011; 365:266–270. <http://www.ncbi.nlm.nih.gov/pubmed/21774716>. [PubMed: 21774716]
32. 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; 10:413–430. [PubMed: 15304142]
33. Cohen J. A power primer. *Psychol. Bull.* 1992; 112:155. [PubMed: 19565683]
34. Westbrook JI, Gosling AS, Coiera E. Do clinicians use online evidence to support patient care? A study of 55,000 clinicians. *J. Am. Med. Inf. Assoc.* 2004; 11:113–120.
35. Mishra R, Bian J, Fiszman M, Weir CR, Jonnalagadda S, Mostafa J, Del Fiol G. Text summarization in the biomedical domain: a systematic review of recent research. *J. Biomed. Inf.* 2014; 52:457–467.
36. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 1989:319–340.
37. Cook DA, Sorensen KJ, Hersh W, Berger RA, Wilkinson JM. Features of effective medical knowledge resources to support point of care learning: a focus group study. *PLoS One.* 2013; 8:e80318. [PubMed: 24282535]

38. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. *ACP J. Club.* 1995; 123:A12–A13. [PubMed: 7582737]
39. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. *J. Biomed. Inf.* 2004; 37:56–76.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Summary points

What was already known on the topic

- Clinicians raise on average one clinical question out of every two patients seen; over half of these questions are left unanswered.
- When used, clinical knowledge resources are able to answer over 90% of clinicians' questions, improving clinicians' performance and patient outcomes.
- Significant barriers still challenge the use of online evidence resources to support clinical decisions at the point of care.

What this study added to our knowledge

- Physicians found the clinical knowledge summary (CKS) prototype investigated in this formative study to be both usable and useful.
- Physicians using the CKS had improved decision quality compared to manual search.
- A decision support tool based on automatic text summarization methods and information visualization principles is a promising alternative to improve clinicians' access to clinical evidence at the point of care.

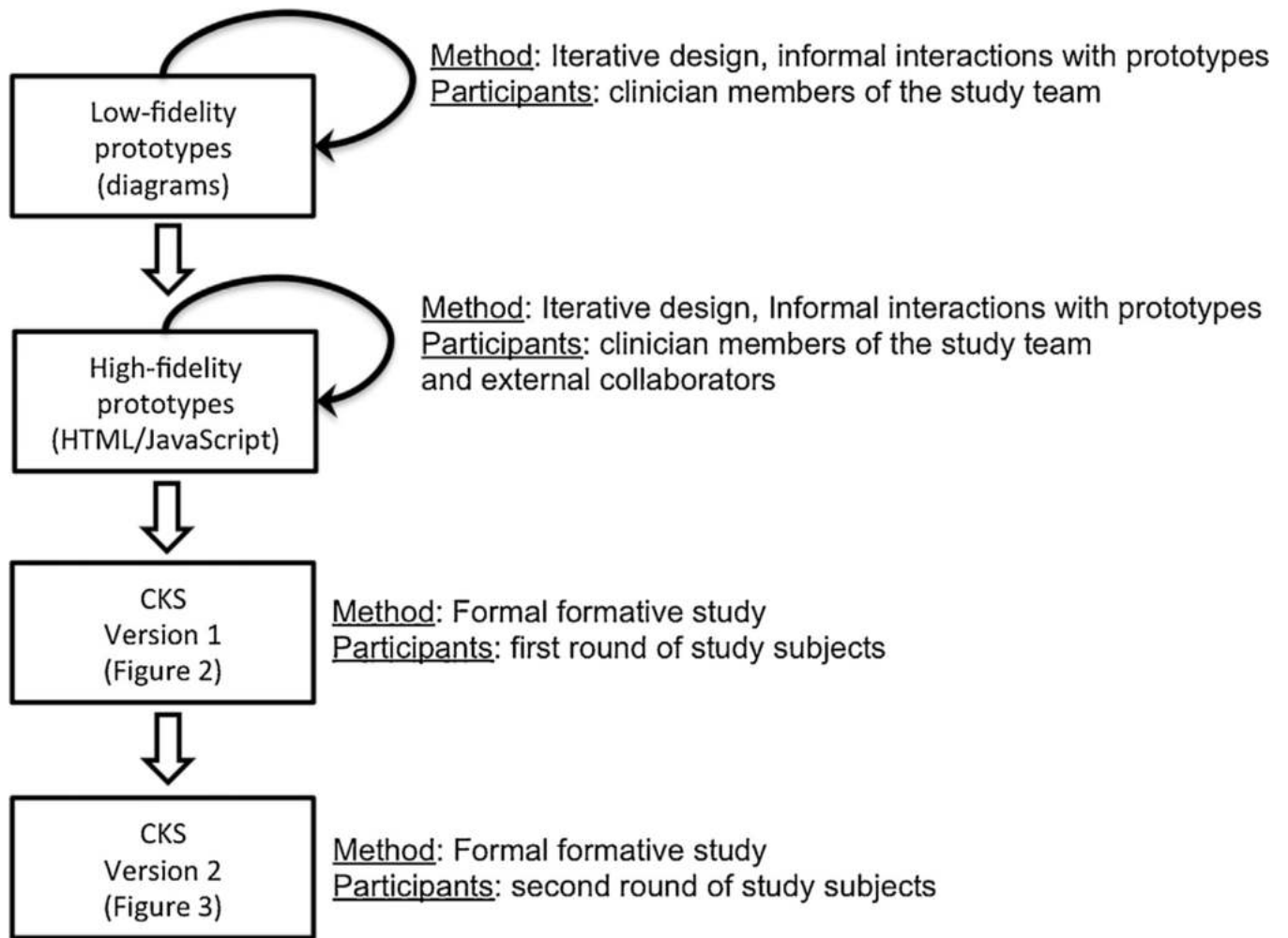


Fig. 1.
Stages of the CKS design process.

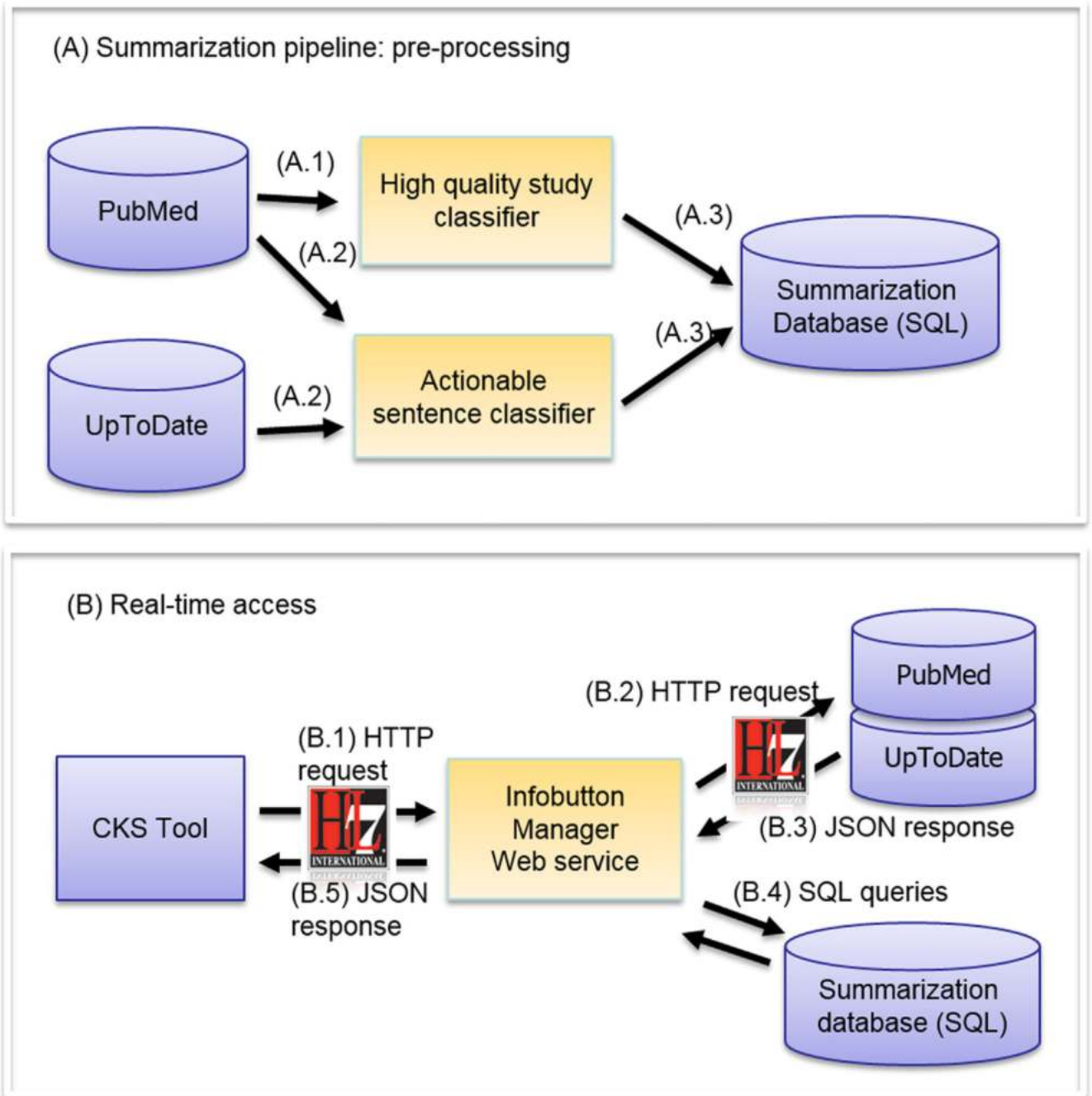


Fig. 2.
CKS architecture.

The screenshot displays the CKS tool interface. On the left, a 'Filters' sidebar lists categories such as 'Biological Response Modifier Therapy (40)', 'Disease-modifying antirheumatic drugs (DMARDs) (292)', 'Glucocorticoids (15)', and 'Tumor necrosis factor (TNF) inhibitors (116)'. The main content area is titled 'UpToDate (276 sentences)' and features a section for 'Initial treatment of moderately to severely active rheumatoid arthritis in adults'. This section includes an introduction, a definition of 'MODERATELY TO SEVERELY ACTIVE RA', pretreatment interventions, and a comparison of MTX versus initial combination therapy. Below this, there are two columns of citations: 'Systematic Reviews (37 citations)' and 'Clinical Trial (110 citations)'. The systematic reviews include titles like 'Effect of biotherapies on fatigue in rheumatoid arthritis: a systematic review of the literature and meta-analysis' and 'Safety of medium- to long-term glucocorticoid therapy in rheumatoid arthritis: a meta-analysis'. The clinical trials include 'Golimumab in patients with active rheumatoid arthritis after treatment with tumour necrosis factor alpha inhibitors (GO-AFTER study): a multicentre, randomised, double-blind, placebo-controlled, phase III trial' and 'Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial'.

Fig. 3. CKS tool after the end of the iterative design phase.

The screenshot displays the CKS tool interface with several callouts highlighting key features:

- Treatment filters:** A sidebar on the left lists various drug classes and their counts, such as 'Biological Response Modifier Therapy (38)', 'Disease-modifying antirheumatic drugs (DMARDs) (280)', and 'Tumor necrosis factor (TNF) inhibitors (110)'.
- UpToDate Excerpts:** The main content area shows a search result for 'Treatment of rheumatoid arthritis resistant to initial DMARD therapy in adults (10 sentences)'. It includes a 'SUMMARY AND RECOMMENDATIONS' section with highlighted text and a 'Link to content within source' callout.
- Sentence counts:** A callout points to the number of sentences in the search result.
- Recommendations from UpToDate:** A callout points to the 'SUMMARY AND RECOMMENDATIONS' section.
- Sentence expansion:** A callout points to a 'more' link that expands a sentence.
- Key treatment terms highlighted:** A callout points to the highlighted drug names 'abatacept', 'tocilizumab', and 'rituximab'.
- Recent high impact systematic reviews:** A callout points to a search result for 'Effect of biotherapies on fatigue in rheumatoid arthritis: a systematic review of the literature and meta-analysis, *Rheumatology (Oxford)*, 2012. Conclusions: Few studies reported the impact of biotherapies on fatigue. The effect of biotherapies on fatigue in RA is small. [more](#)'.
- Recent high impact RCTs:** A callout points to a search result for 'Treatment impact on estimated medical expenditure and job loss likelihood in rheumatoid arthritis: re-examining quality of life outcomes from a randomized placebo-controlled clinical trial with abatacept, *Rheumatology (Oxford)*, 2008. Conclusions: QoL changes provided greater reduction in medical expenditure and likelihood of an inability to work. The strong effect sizes obtained in this analysis suggest that the results are clinically important. However, because of the many variables, results should also be interpreted with caution. Further studies regarding using estimated outcomes are needed. [more](#)'.
- Article title, journal, publication date, funding source, and sample size:** A callout points to the structured data provided for each search result.
- Article conclusions:** A callout points to the 'Conclusions' section of a search result.
- Click to display article results:** A callout points to a 'more' link that expands the search result.

Fig. 4. CKS tool after refinement based on exposure to the first set of participants in the formative evaluation.

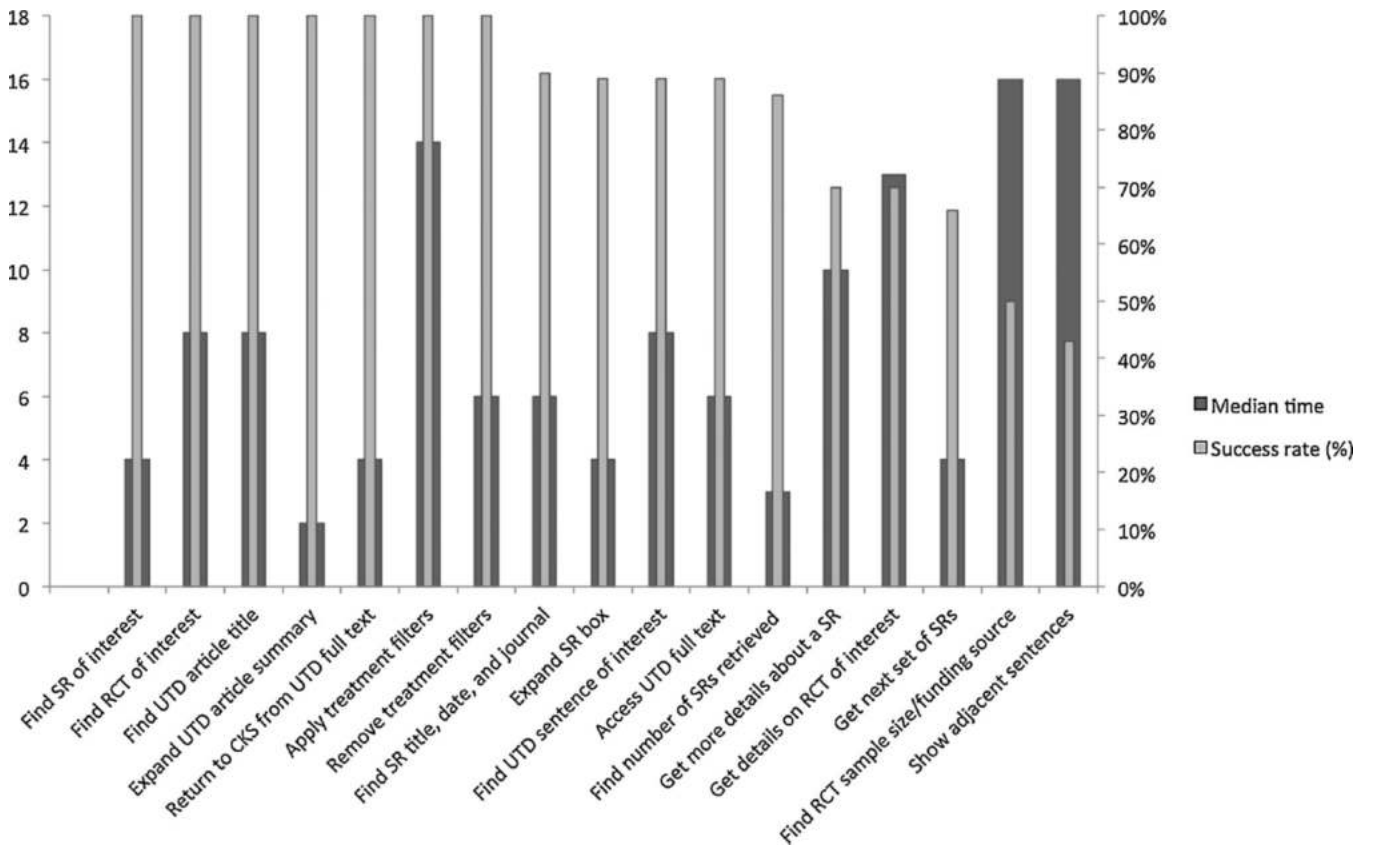


Fig. 5. Usability task time to completion (in seconds) and completion success rate.

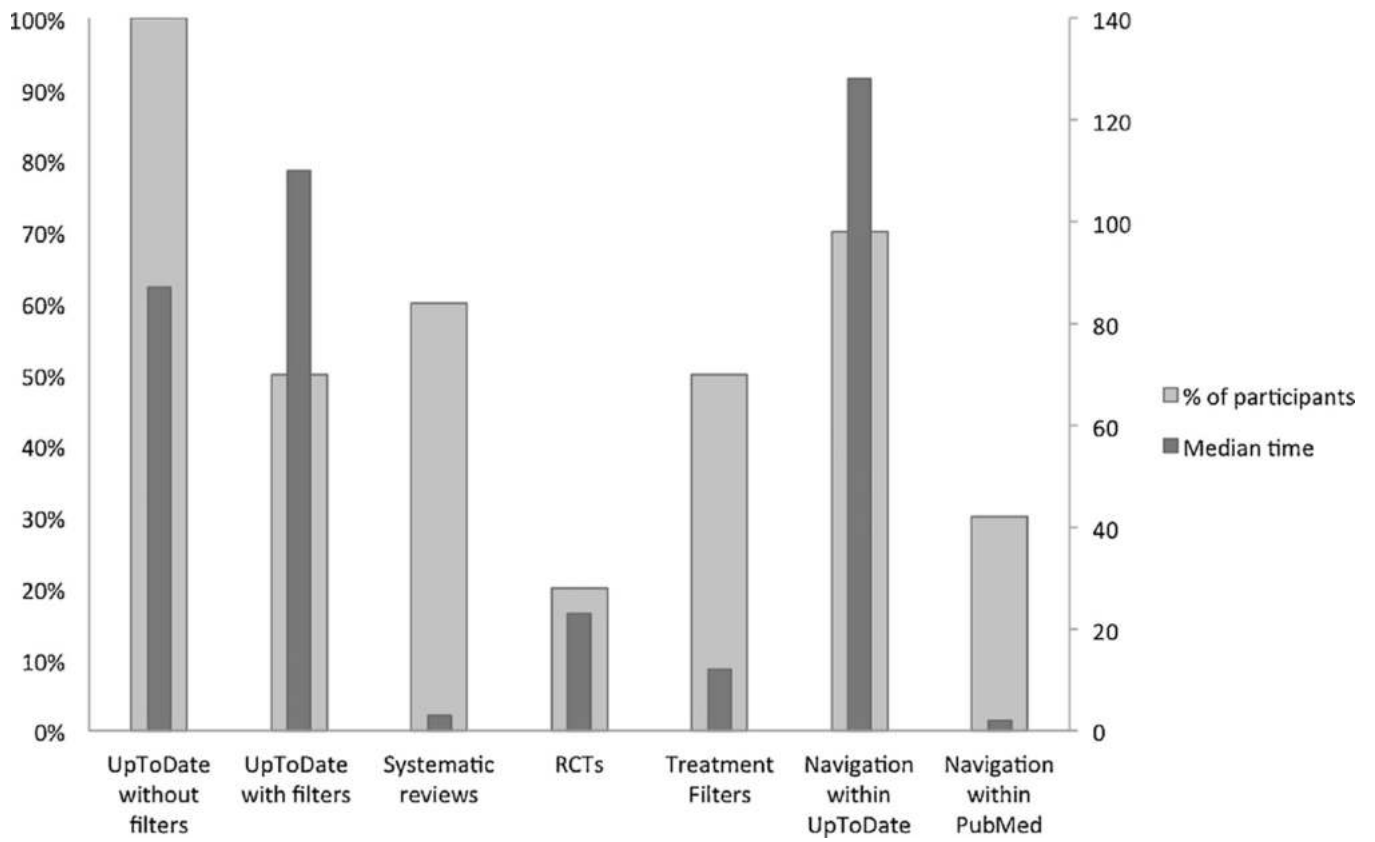


Fig. 6. Percentage of users who used each CKS feature and mean time spent on each CKS feature.

Table 1

Usability task time to completion (in seconds) and completion success rate.

Usability task	Average time (s)	Median time [range]	Success rate
Find SR of interest	12 ± 24	4 [1–80]	100%
Find RCT of interest	26 ± 32	8 [2–84]	100%
Find UTD article title	8 ± 4	8 [3–16]	100%
Expand UTD article summary	4 ± 3	2 [1–10]	100%
Return to CKS from UTD full text	8 ± 9	4 [1–29]	100%
Apply treatment filters	17 ± 13	14 [4–47]	100%
Remove treatment filters	8 ± 4	6 [5–14]	100%
Find SR title, publication date, and journal	27 ± 38	6 [1–105]	90%
Expand SR box	5 ± 3	4 [1–10]	89%
Focus on one UTD sentence of interest	13 ± 12	8 [5–38]	89%
Access UTD full text	16 ± 25	6 [3–80]	89%
Find number of SRs retrieved	3 ± 3	3 [1–8]	86%
Get more details about a SR	16 ± 15	10 [3–44]	70%
Get more details about an RCT of interest	19 ± 14	13 [3–37]	70%
Get next set of SRs	5 ± 6	4 [3–13]	66%
Find RCT sample size/funding source	22 ± 25	16 [1–96]	50%
Display adjacent sentences	20 ± 12	16 [9–41]	43%

Table 2

Time (in seconds) spent on and percentage of users who used each CKS feature.

CKS feature	Average time (s)	Median time [range]	% of participants who accessed the feature
UpToDate excerpts without filters	91 ± 70	87 [4–117]	100%
Navigation within UpToDate	144 ± 74	128 [2–242]	70%
Systematic reviews	26 ± 45	20 [6–118]	60%
UpToDate excerpts with filters	115 ± 81	110 [8–167]	50%
Treatment filters	15 ± 11	12 [4–43]	50%
Navigation within PubMed	21 ± 34	117 [58–175]	30%
RCTs	23 ± 17	23 [11–35]	20%

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Participant satisfaction with CKS features.

Feature	Average rating (1–5) [range]
Link to source article in UpToDate	4.6 [3–5]
Treatment filters	4.1 [3–5]
Expand adjacent sentences	3.9 [3–5]
Article titles	3.7 [1–5]
Maximize button	3.1 [1–4]
Link to source abstract in PubMed	2.9 [1–5]
Study funding source	2.4 [1–4]
Study sample size	2.3 [1–4]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Partial correlations (controlling for vignette experience) between *perceived decision quality* and explanatory variables of *improved recall*, *surprise*, *effort*, *frustration* and *likelihood to refer*.

	Perceived decision quality	
	Manual search	CKS
Improved recall	0.32	0.47
Surprise	-0.15	-0.19
Effort	-0.32	0.21
Frustration	0.60 ^a	0.63 ^a
Likelihood to refer	0.87 ^a	-0.14

^a Statistically significant.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript