



Physicians' perception of alternative displays of clinical research evidence for clinical decision support – A study with case vignettes



Stacey L. Slager^a, Charlene R. Weir^{a,b}, Heejun Kim^c, Javed Mostafa^c, Guilherme Del Fiol^{a,*}

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

^b George E Whalen VA Medical Center, Salt Lake City, UT, USA

^c School of Information and Library Science, University of North Carolina, Chapel Hill, NC, USA

ARTICLE INFO

Article history:

Received 4 February 2016

Revised 12 September 2016

Accepted 11 January 2017

Available online 13 January 2017

Keywords:

Information display

Information seeking

Clinical decision making

Clinician information needs

Information foraging theory

ABSTRACT

Objective: To design alternate information displays that present summaries of clinical trial results to clinicians to support decision-making; and to compare the displays according to efficacy and acceptability.

Methods: A 6-between (information display presentation order) by 3-within (display type) factorial design. Two alternate displays were designed based on Information Foraging theory: a narrative summary that reduces the content to a few sentences; and a table format that structures the display according to the PICO (Population, Intervention, Comparison, Outcome) framework. The designs were compared with the summary display format available in PubMed. Physicians were asked to review five clinical studies retrieved for a case vignette; and were presented with the three display formats. Participants were asked to rate their experience with each of the information displays according to a Likert scale questionnaire.

Results: Twenty physicians completed the study. Overall, participants rated the table display more highly than either the text summary or PubMed's summary format (5.9 vs. 5.4 vs. 3.9 on a scale between 1 [strongly disagree] and 7 [strongly agree]). Usefulness ratings of seven pieces of information, i.e. *patient population*, *patient age range*, *sample size*, *study arm*, *primary outcome*, *results of primary outcome*, and *conclusion*, were high (average across all items = 4.71 on a 1 to 5 scale, with 1 = not at all useful and 5 = very useful). *Study arm*, *primary outcome*, and *conclusion* scored the highest (4.9, 4.85, and 4.85 respectively). Participants suggested additional details such as rate of adverse effects.

Conclusion: The table format reduced physicians' perceived cognitive effort when quickly reviewing clinical trial information and was more favorably received by physicians than the narrative summary or PubMed's summary format display.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

One of the goals of biomedical research is to provide clinicians with the best evidence possible to support patient care decisions. Despite best efforts, tools that provide optimal access to evidence-based research reports in the primary literature for clinical decision support are still lacking [1–3]. This is important, because clinicians often raise clinical questions in the course of patient care that can be answered by online clinical evidence resources such as Medline [4]. When faced with a non-routine clinical problem, having clinicians consult the primary literature for the best available evidence might seem ideal, yet the size, evolving

nature, and complexity of the primary literature make it challenging for consumption at the point of decision-making. As a result, physicians prefer to consult alternate sources such as books, colleagues, guidelines, and synthesized sources of evidence such as UpToDate[®] and Dynamed[®] [5,6].

When faced with a clinical question, clinicians need to scan a potentially large number of citations to find those that are relevant for a particular patient [7–9]. This process is cognitively intense and often prohibitive in busy clinical workflows. Several tools have been investigated to improve clinicians' access to the primary literature, but most of these efforts focused on improving the search process to find high quality articles [8–14]. Less effort has been dedicated to tools that improve clinicians' ability to quickly scan citations to identify those that apply to a specific patient and information needs.

* Corresponding author at: Department of Biomedical Informatics, University of Utah, 421 Wakara Way, Salt Lake City, UT 84108, USA.

E-mail address: guilherme.delfiol@utah.edu (G. Del Fiol).

In the present study, we designed alternate information displays for presenting the results of clinical studies to clinicians. The alternate displays were designed according to principles derived from Information Foraging theory [15]. Applying Information Foraging theory, we speculate that scanning for patient-specific studies can be streamlined with structured data visualization that provides cues to help information seekers to quickly identify relevant information. We hypothesized that, compared to traditional narrative formats, an information display that minimizes visual clutter and uses a semi-structured format improves clinicians' perceived efficiency and efficacy when scanning for relevant studies for clinical case vignettes.

2. Materials and methods

2.1. Overview

This study used a within-subject design where each participant was exposed to three types of presentation formats associated with one out of three possible case vignettes on the following topics: (i) efficacy and safety of vernakalant for cardioversion in atrial fibrillation (Fig. 1); (ii) benefits and safety of statins for a patient with heart failure, hypertension and previous myocardial infarction; and (iii) evidence on the use of metformin combined with a dipeptidyl peptidase-4 (DPP-4) inhibitor for a patient with uncontrolled diabetes mellitus. Each vignette consisted of a brief case narrative and five reports of clinical trials potentially relevant to the decision task presented in the vignette. The vignettes described simple clinical cases that were meant to focus the user's search to determine the relevancy of the search results, rather than for problem solving. The full vignette descriptions are available in the [online supplement](#).

2.2. Information display design

We compared three information displays that summarize clinical trial reports. The first display was the standard presentation of search results available in *PubMed* (Fig. 2). The second display consisted of a synthesized *narrative summary* with key information from the abstract including title, journal, publication year, and a few sentences about the study conclusions (Fig. 3). The third display consisted of a semi-structured *table format* according to elements in the *PICO framework* (Population, Intervention, Comparison, and Outcome) (Fig. 4) [16].

2.2.1. Design rationale

Our design decisions were guided by *Information Foraging theory* [15] and the *PICO framework* [16]. Information Foraging proposes that humans forage for information in *information patches* similar to how animals forage for food. In the information foraging process, information seekers constantly estimate the cost and benefit of foraging from a specific *information patch*. Our approach focused on helping information seekers produce cost-benefit estimates more reliably and efficiently so they can quickly identify

the most profitable *information patches*. Specifically, we adopted the following strategies: (1) maximize cues (i.e., *information scent*) that help seekers to quickly identify relevant information; and (2) apply *information patch enrichment*, i.e. ensure that information patches have the highest possible concentration of useful information.

The PICO framework emerged in the mid-1990s as a means for physicians and clinical researchers to frame their clinical questions [17,18]. Since then, it has been widely adopted for evidence-based inquiry, especially in the development of systematic reviews and clinical practice guidelines [19–24]. Clinicians have also been recommended to structure their patient-specific clinical questions according to PICO [18,25–29]. Finally, information retrieval tools that structure queries according to the PICO format have shown promising results [30–32]. However, searches formulated in the PICO format in current search engines retrieve publications that are then displayed in a format that follows the traditional abstracts in biomedical journals [16]. This leads to a mismatch between the clinicians' mental model that is reflected in the structure of the search and the format of the retrieved information, compromising *information scent*. In addition, several pieces of information in a scientific abstract may not be relevant to clinicians' decision-making, compromising *information patch enrichment*.

Our design goals were to maximize *information scent* by matching the structure of the information display to clinicians' information foraging mental model; and by including only the information that is necessary for judging the relevance of clinical trial publications for a specific patient (the Population, Intervention, and Comparison in the PICO format) and interpreting the gist of the findings of those trials, therefore maximizing *patch enrichment*. The use of table format information displays have been shown to be successful in other studies and supports our hypothesis [33].

2.3. Preparation of display formats

Using iterative design processes, members of our research team and target users provided feedback on a series of hypertext markup language (HTML) prototypes. The motivation for the design alternatives was to present only the information necessary to help clinicians judge the relevancy of a study to a specific patient and to enhance the ability of readers to grasp the gist of the study findings rapidly. For each vignette, we selected publications resulting from five recent clinical trials that were manually searched using *PubMed's Clinical Queries filter*. The key information about the five clinical trials was then manually formatted in HTML according to the two alternate displays (Figs. 3 and 4).

2.4. Participants/recruitment

Participants for this study were a convenience sample of 20 physicians who were invited by e-mail to join our study. The first participants were known to members of our team. From that initial sample we used snowball sampling, where participants recruited others like themselves to participate in the study, to achieve a lar-

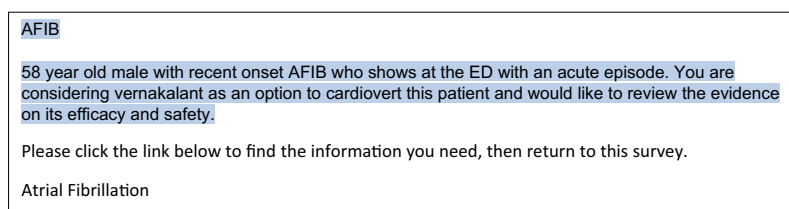


Fig. 1. Atrial fibrillation vignette.

Results: 5

- [Effect of dietary fish oil on atrial fibrillation after cardiac surgery.](#)
 1. Farquharson AL, Metcalf RG, Sanders P, Stuklis R, Edwards JR, Gibson RA, Cleland LG, Sullivan TR, James MJ, Young GD.
 Am J Cardiol. 2011 Sep 15;108(8):851-6. doi: 10.1016/j.amjcard.2011.04.036. Epub 2011 Jul 15.
 PMID: 21762871
[Similar articles](#)
- [A randomized active-controlled study comparing the efficacy and safety of vernakalant to amiodarone in recent-onset atrial fibrillation.](#)
 2. Camm AJ, Capucci A, Hohnloser SH, Torp-Pedersen C, Van Gelder IC, Mangal B, Beatch G; AVRO Investigators.
 J Am Coll Cardiol. 2011 Jan 18;57(3):313-21. doi: 10.1016/j.jacc.2010.07.046.
 PMID: 21232669 **Free Article**
[Similar articles](#)
- [Vernakalant hydrochloride: A novel atrial-selective agent for the cardioversion of recent-onset atrial fibrillation in the emergency department.](#)
 3. Stiell IG, Dickinson G, Butterfield NN, Clement CM, Perry JJ, Vaillancourt C, Calder LA.
 Acad Emerg Med. 2010 Nov;17(11):1175-82. doi: 10.1111/j.1553-2712.2010.00915.x. Erratum in: Acad Emerg Med. 2011 Feb;18(2):224.
 PMID: 21175515
[Similar articles](#)
- [Usefulness of vernakalant hydrochloride injection for rapid conversion of atrial fibrillation.](#)
 4. Pratt CM, Roy D, Torp-Pedersen C, Wyse DG, Toft E, Juul-Møller S, Retyk E, Drenning DH; Atrial Arrhythmia Conversion Trial (ACT-III) Investigators.
 Am J Cardiol. 2010 Nov 1;106(9):1277-83. doi: 10.1016/j.amjcard.2010.06.054.
 PMID: 21029824
[Similar articles](#)
- [Pharmacokinetics of novel atrial-selective antiarrhythmic agent vernakalant hydrochloride injection \(RSD1235\): influence of CYP2D6 expression and other factors.](#)
 5. Mao ZL, Wheeler JJ, Clohs L, Beatch GN, Keims J.
 J Clin Pharmacol. 2009 Jan;49(1):17-29. doi: 10.1177/0091270008325148. Epub 2008 Oct 16.
 PMID: 18927241
[Similar articles](#)

Fig. 2. PubMed summary display format for atrial fibrillation. Participants had to click on each article title to read the study abstract.

ger sample size. The recruitment targeted practicing primary care and internal medicine physicians with a wide range of clinical experience, from second year residents to practitioners with over 25 years of patient care experience.

2.5. Procedure

The e-mail sent to potential participants included a link to an online survey, from which they could launch the case vignettes along with the three information displays. Participants completed the assessment independently and at their own leisure with no time limit imposed. Participants were blinded to the study goals and hypotheses. For each participant, three vignettes were randomly assigned to one of the display types, so that each participant viewed search results for all three vignettes, each result displayed in a different display format. Randomization was implemented with a simple computer-based randomization procedure. There were six possible vignette-information display pairings. Each pairing was assigned a range within 0 and 1. At the time of enrollment of each participant, a computer algorithm generated a random number from 0 to 1, which determined the vignette-display pairing for the participant.

Participants did not receive training on any of the information displays. Information displays were presented in random order for each participant to minimize order effect, with no washout period between vignette presentations. After reading each vignette, participants were asked to scan the summary information about the clinical trials, find the publications that were relevant to the

vignette, and interpret the key results. At the end of each vignette and information visualization, participants were asked to complete a brief questionnaire in the online survey and data collection tool Research Electronic Data Capture REDCap [34]. The study was approved by the University of Utah Institutional Review Board.

2.6. Measures

We developed an 11-item questionnaire that included 4 questions obtained from the System Usability Scale (SUS) [35]; 6 questions that measured self-perceived ability to understand the meaning of user interface features, quickly scan and judge study relevancy to the vignette, and interpret the studies that were presented; and one question about satisfaction. Questions were framed according to a Likert scale (1 = strongly disagree; 7 = strongly agree). The complete questionnaire is available in the [online supplement](#).

Participants were asked an additional set of questions about the tabular display format to rate the usefulness of the set of display features we had chosen to present (Table 2) according to a 5-point Likert scale (1 = not at all useful, 5 = very useful) with an option to add open-ended comments.

2.7. Data analysis

The data analysis assessed the differential impact of the 3 different displays on perceptions of decision quality, usability and satisfaction using a generalized linear model with repeated mea-

Clinical Trial

[Vernakalant hydrochloride: A novel atrial-selective agent for the cardioversion of recent-onset atrial fibrillation in the emergency department.](#) *Acad Emerg Med.* 2010. [Industry funding]. [n=356].

Conclusions: Vernakalant rapidly converted recent-onset AF to sinus rhythm in over half of patients, was well tolerated, and has the potential to offer an important therapeutic option for rhythm control of recent-onset AF in the ED. [more](#)

[Pharmacokinetics of novel atrial-selective antiarrhythmic agent vernakalant hydrochloride injection \(RSD1235\): influence of CYP2D6 expression and other factors.](#) *J Clin Pharmacol.* 2009. **Conclusions:**

The impact of CYP2D6 metabolizer status on vernakalant exposure was explored in patients with atrial fibrillation or atrial flutter who received a therapeutic regimen (3 mg/kg initially via 10-minute intravenous infusion followed by a second 2 mg/kg 10-minute infusion if atrial fibrillation persisted after a 15-minute observation period). These results suggest that an assessment of CYP2D6 expression may not be needed when vernakalant is administered acutely and intravenously to patients with atrial fibrillation.

[Usefulness of vernakalant hydrochloride injection for rapid conversion of atrial fibrillation.](#) *Am. J. Cardiol.* 2010. **Conclusions:**

The primary end point was conversion of AF to SR for ≥ 1 minute within 90 minutes of the start of the drug infusion in the short-duration AF group. Of the 86 patients receiving vernakalant in the short-duration AF group, 44 (51.2%) demonstrated conversion to SR compared to 3 (3.6%) of the 84 in the placebo group ($p < 0.0001$). The median interval to conversion of short-duration AF to SR in the responders given vernakalant was 8 minutes. Of the entire AF population (short- and long-duration AF), 47 (39.8%) of the 118 vernakalant patients experienced conversion of AF to SR compared to 4 (3.3%) of the 121 placebo patients ($p < 0.0001$).

[Effect of dietary fish oil on atrial fibrillation after cardiac surgery.](#) *Am. J. Cardiol.* 2011. **Conclusions:**

There was a nonstatistically significant delay in time to onset of AF in the fish oil group (hazard ratio 0.66, 95% CI 0.43 to 1.01). There was a significant decrease in mean length of stay in the intensive care unit in the fish oil group (ratio of means 0.71, 95% CI 0.56 to 0.90). In conclusion, in a mixed cardiac surgery population, supplementation with dietary fish oil did not result in a significant decrease in the incidence of postsurgical AF. However, there was a significant decrease in time spent in the intensive care unit.

[A randomized active-controlled study comparing the efficacy and safety of vernakalant to amiodarone in recent-onset atrial fibrillation.](#) *J. Am. Coll. Cardiol.* 2011. [Industry funding]. [n=254]. **Conclusions:**

Vernakalant demonstrated efficacy superior to amiodarone for acute conversion of recent-onset AF. Both vernakalant and amiodarone were safe and well tolerated in this study. (A Phase III Superiority Study of Vernakalant vs Amiodarone in Subjects With Recent Onset Atrial Fibrillation [AVRO];

Fig. 3. Narrative summary display for a case vignette on atrial fibrillation.

Vernakalant hydrochloride: A novel atrial-selective agent for the cardioversion of recent-onset atrial fibrillation in the emergency department. <i>Acad Emerg Med.</i> 2010.			
Population	Primary Outcome	Study arm/results	Conclusion
Adults with recent-onset AF (> 3 to <= 48 hours) presenting to the ED (n=290)	Conversion to sinus rhythm within 90 minutes	59.4% vernakalant vs. 4.9% placebo	Vernakalant rapidly converted recent-onset AF to sinus rhythm in over half of patients, was well tolerated, and has the potential to offer an important therapeutic option for rhythm control of recent-onset AF in the ED.
Pharmacokinetics of novel atrial-selective antiarrhythmic agent vernakalant hydrochloride injection (RSD1235): influence of CYP2D6 expression and other factors. <i>J Clin Pharmacol.</i> 2009.			
Healthy volunteers and patients with atrial fibrillation or atrial flutter	Maximum plasma concentration	Little difference in vernakalant maximum plasma concentration or area under the plasma concentration-time curve from the start of the first infusion to 90 minutes between CYP2D6 poor metabolizers and extensive metabolizers.	Gender, age, and renal function did not have a clinically significant influence on the pharmacokinetics of vernakalant. An assessment of CYP2D6 expression may not be needed when vernakalant is administered acutely and intravenously to patients with atrial fibrillation.
Usefulness of vernakalant hydrochloride injection for rapid conversion of atrial fibrillation. <i>Am J Cardiol.</i> 2010.			
AF or atrial flutter with arrhythmia duration of >3 hours to <=45 days (n=276)	Conversion of AF to SR	39.8% vernakalant vs. 3.3% placebo ($p < 0.0001$) Transient dysgeusia and sneezing were the most common adverse events.	Vernakalant demonstrated a rapid and high rate of conversion for short-duration AF and was well tolerated.
Effect of dietary fish oil on atrial fibrillation after cardiac surgery. <i>Am J Cardiol.</i> 2011.			
Postsurgical AF after CABG and valve procedures (n=200)	Incidence of AF in the first 6 days after surgery	47 of 97 in the control group vs. 36 of 97 in the fish oil group developed AF (OR 0.63, 95% CI 0.35 to 1.11)	Supplementation with dietary fish oil did not result in a significant decrease in the incidence of postsurgical AF. However, there was a significant decrease in time spent in the intensive care unit.
A randomized active-controlled study comparing the efficacy and safety of vernakalant to amiodarone in recent-onset atrial fibrillation. <i>J Am Coll Cardiol.</i> 2011.			
Adult patients with AF (3 to 48 h duration) eligible for cardioversion (n=254)	Conversion from AF to sinus rhythm within the first 90 min	60 of 116 (51.7%) vernakalant vs. 6 of 116 (5.2%) amiodarone ($p < 0.0001$).	Vernakalant demonstrated efficacy superior to amiodarone for acute conversion of recent-onset AF. Both vernakalant and amiodarone were safe and well tolerated in this study.

Fig. 4. PICO tabular display format for a case vignette on atrial fibrillation.

tures factorial analysis with design type as a within-subject component and sequence as a between-subject variable. Mean differences were assessed using a generalized linear model repeated measures analysis with design type as the within-subject variable and sequence as the between-subject variable. The statistical power of the sample size for all study comparisons was above 80%.

3. Results

3.1. Mean comparison across groups

Across all of the single-item questions, the PICO presentation format received significantly higher ratings after controlling for sequence in all 11 variables (Table 1). In all of the single-item analyses, the effect of sequence was non-significant. The Benjamini-Hochberg method was used to control for multiple comparisons [36] and all but one (Q10) remained significant. All post hoc Tukey comparisons were not significant. A power analysis was done assessing a within-subject ANOVA for 20 and 2 trials and a large effect size estimate of 0.70 with a resultant power estimate of 0.844 and a non-centrality parameter of 3.13.

3.2. Usefulness ratings of specific PICO display components

The average ratings for the different types of information presented in the PICO tabular display was between 4.4 and 4.9 out of 5 (Table 2).

3.3. Open-ended comments

The open-ended comments in general favored the PICO display and included suggestions for improvement (Table 3).

4. Discussion

Clinicians must engage in substantial efforts to address their frequent scientific questions and many go unanswered. Although substantial work has been dedicated to improving the biomedical literature search process [8,9,12,37], less has been done to investigate optimal displays for summarizing individual studies, especially in the context of clinical decision support [38]. This study was designed to address this question by assessing the perceived efficiency and efficacy of 3 display formats for presenting a summary of clinical trial reports to clinicians. Participants significantly rated either the *PICO tabular display* or the narrative display higher than the PubMed display in at least 10 of the 11 criteria. Of the two non-PubMed displays, *PICO tabular displays* were always rated

Table 2

Usefulness of different types of information presented in the PICO tabular display (1 = not at all useful, 5 = very useful).

Item	Mean ± standard deviation
Patient population	4.8 ± 0.4
Patient age	4.6 ± 1.0
Sample size	4.4 ± 1.1
Study arm	4.7 ± 0.6
Primary outcome	4.9 ± 0.3
Primary outcome results	4.9 ± 0.5
Conclusion	4.9 ± 0.4

higher, but post hoc comparisons did not show a difference, although the pattern is suggestive. Our results suggest that a tabular format is helpful for communicating useful and clinically actionable information when compared to a traditional PubMed format. One explanation for our findings may be that a structure that distills the main components of a clinical trial may require less cognitive load than the traditional PubMed narrative summary, no matter how concise. Our study suggests that further work on creating tabular displays would be a promising alternative to help clinicians find answers to their patient-specific clinical questions in the primary literature and would warrant further investigation, both in terms of identifying optimal displays and developing algorithms that can automate tabular displays [15]. Other fields have found similar positive results for tabular displays [33].

The *PICO tabular* design is particularly supported by empirical and theoretical principles. The PICO framework has been widely promoted as a mechanism to help clinicians formulate better clinical questions [16]. In addition, participants in a study of summarized systematic reviews suggested that the most useful data in a clinical setting would include “population, setting, intervention and control group” [39]. The *PICO table* was designed to optimize information seeking by following a consistent structure that matches clinicians' information seeking mental model, therefore increasing *information scent* [15]. Although the post hoc comparisons were non-significant, the consistent pattern of higher scores for the PICO format or the tighter narrative supports the notion that this framework might be successful in a larger trial.

Participants rated the *text summary* and the *PICO tables* higher than the *PubMed summary displays* in terms of scanning the information quickly, judging the relevancy of studies for a particular patient, and overall usability. Because previous research has shown that clinicians are unlikely to devote more than two to three minutes for such an information search [4], supporting efficient information scanning and rapid relevance judgment was one of the most important design considerations in the alternate displays.

Table 1

Physicians' ratings of the three information displays. All ratings use a Likert scale 1 (strongly disagree) to 7 (strongly agree).

Criteria	PubMed	Narrative summary	PICO table	F score	P value /adjusted (0.05)
	Mean ± standard deviation				
Q1. It was easy to understand the meaning of the information presented.	4.4 ± 1.7	5.8 ± 1.2	5.9 ± 1.2	6.0	0.008/0.05
Q2. I think that I would like to use this product frequently.	3.6 ± 1.76	5.6 ± 1.3	5.9 ± 1.3	16.8	<0.001/0.05
Q3. I was able to scan the information quickly.	3.5 ± 1.8	5.8 ± 1.2	6.0 ± 1.3	18.6	<0.001/0.05
Q4. I was able to quickly determine relevance of the study for my patient.	4.2 ± 1.7	5.3 ± 1.3	6.1 ± 1.1	6.0	0.008/0.05
Q5. It was easy for me to locate the information about the study that I needed.	3.0 ± 1.2	4.6 ± 1.7	5.2 ± 1.6	8.5	0.002/0.05
Q6. I thought that this product was easy to use.	4.1 ± 1.6	6.0 ± 1.4	6.2 ± 1.2	15.9	<0.001/0.05
Q7. I was able to find the pieces of information I needed to understand the study.	4.2 ± 1.8	5.6 ± 1.2	6.2 ± 1.0	19.7	<0.001/0.05
Q8. I would imagine that most people would learn to use this product quickly.	4.9 ± 1.5	5.6 ± 1.5	6.1 ± 1.2	8.0	0.002/0.05
Q9. I was able to quickly grasp the gist of the paper's findings.	3.7 ± 1.9	5.6 ± 1.1	6.1 ± 1.0	13.1	<0.001/0.05
Q10. I found this product very awkward to use. (reversed question)	3.8 ± 1.8	5.4 ± 1.5	5.7 ± 1.4	4.4	0.02/NS
Q11. In general, I am satisfied with the presentation (i.e., format of the display) of the information.	3.7 ± 1.6	5.4 ± 1.3	5.7 ± 1.3	13.9	<0.001/0.05
Average and standard deviation of all items (after reversing the scale of negative criteria)	3.9 ± 0.5	5.4 ± 0.4	5.9 ± 0.3	N/A	N/A

Table 3
Open-ended comments on the PICO display.

In the tabular display, we have included information on patient population, age, sample size, study arms, outcomes, and conclusions. Is there enough information here for you to be able to make a clinical decision about the patient?

Yes! It was great to click on the study and see the full abstract and links to other studies. It was so much faster to have the studies broken down and the information organized. This is generally enough.

Yes. 'Sample size' was listed in #25, but didn't see it appear consistently, making it hard to know if the trial was powered sufficiently. Indication of statistical significance would clearly be useful.

Generally yes.

The tabular form when introduced secondarily, at first overloads because of the complexity of the layout but quickly overcomes this by its utility. This presentation provides structure to the text from the articles without removing the semantic benefit of transmitting information through prose and allows the reader to easily compare study characteristics.

There was a lot of information at first. After using it a few times it would probably be quite a bit easier to understand it quickly.

Other types of information that could be included and display suggestions.

Race, was it an industry funded study? A link to editorials on the paper if there are any.

Study limitations

It would be very helpful to know if the study can be applied to the patient for whom a clinical decision is being made. For example, in the geriatrics clinic, it is essential to know whether older adults were studied, and if so, if there were adequate numbers of older adults in the trial to make a meaningful conclusion. I would add that the study arm/results column was somewhat difficult to read, as I think it is difficult to condense this type of detailed information into a small space.

Need to read abstract for full context. For example, abstract for first study included this useful sentence: Rosuvastatin 10 mg daily did not affect clinical outcomes in patients with chronic heart failure of any cause, in whom the drug was safe.

Type of study (randomized, placebo-controlled, case control, etc.), follow-up period, lost to follow-up rate, measures that would suggest effectiveness of randomization (or other variables controlled for)

It would be nice to color code the results by positive effect, negative effect or no measurable effect.

Another important feature was optimizing the display to improve the user's ability to quickly understand the gist of study findings. Overall, our findings are congruent with the Technology Acceptance Model (TAM) [40], which postulates that perceived ease of use and perceived usefulness are predictors of actual adoption of a particular technology.

Based on user suggestions, other items could be included in the table such as race, funding source, study type, and adverse effects. However, additional information could also clutter the display and compromise efficiency and cognitive effort. A potential alternative is to add interactive functions to the PICO display so that only the essential information for scanning and relevance judging is displayed up-front and additional details are provided on-demand upon user request.

4.1. Limitations

This study was focused on assessing participants' self-perceived ability to scan, judge relevancy, and interpret the results of clinical studies. We have not investigated the effect of the different information displays on physicians' actual decision making subjected to typical clinical setting constraints, such as time pressure and interruptions. Yet, the perceived improvements of the PICO display in items related to efficiency and cognitive effort suggest clinician adoption. Future studies could present clinical vignettes that request clinicians to make patient care decisions, simulating an environment with time pressure and interruptions. We limited the displays to five publications to minimize participants' time. Actual searches are likely to retrieve a much larger number of studies. Yet, PubMed displays by default up to 20 publications per page and clinicians are unlikely to scan a large number of studies in the context of patient care [4]. The three medical conditions we used for the case vignettes are fairly common conditions. It is possible that other design features may be needed to help with clinical scenarios that are less familiar to clinicians. All participants were previously familiar with the PubMed format and a subset of the study subjects ($n = 4$) had participated in usability testing of the narrative summary display. Familiarity with those displays may have influenced participants' ratings in comparison with the PICO table in ways that we were unable to measure. Last, we have not investigated the effect of clinical experience and specialty on the study outcomes, which limits the generalizability of findings with different populations. Yet, we expect that the randomization of partici-

pants across the six order pairs and the within-subject design minimizes the impact of individual differences as an explanation of the findings.

4.2. Future work

Future investigation is needed to identify optimal displays based on the PICO table, developing algorithms that can automate tabular displays, and investigate the use of tabular displays in real clinical settings. While the information items provided in the PICO table were all rated highly in terms of usefulness, some participants requested that other pieces of information be included, such as adverse events. Researchers have investigated mechanisms to automatically extract key pieces of information from clinical trials using methods such as information extraction from full-text articles [41–43]. Our research group is currently investigating the feasibility of extracting PICO elements from ClinicalTrials.gov. We are also investigating alternative displays that add interactive functionality to the PICO table.

5. Conclusion

Overall, a tabular information display structured according to the PICO framework appears to be more useful to clinicians than a traditional (i.e., PubMed) narrative format, which aims to serve general audiences and investigators. Integration of primary literature resources with EHR systems for clinical decision support could be improved with a tabular PICO display and warrant further investigation. While we have not investigated the effect of the different displays on clinicians' decisions and clinical outcomes, superiority of the PICO table display in efficiency and effort favor clinician adoption of such a display format and warrant further research.

Conflict of interest

None.

Funding and role of funding source

This project was supported by grant 1R01LM011416-01 from the National Library of Medicine. The funders had no role in the

study design, collection, analysis, and interpretation of the data, the writing of the report or the decision to submit the article for publication.

Disclaimer

The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government or any of the affiliated institutions.

Acknowledgements

The authors would like to acknowledge Innovation Librarian Tallie Casucci from the University of Utah Eccles Library for her assistance with the literature search.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jbi.2017.01.007>.

References

- [1] L.H. Reeve, H. Han, A.D. Brooks, Biomedical text summarisation using concept chains, *Int J Data Min Bioinform* 1 (4) (2007) 389–407.
- [2] D.A. Cook, K.J. Sorensen, W. Hersh, R.A. Berger, J.M. Wilkinson, Features of effective medical knowledge resources to support point of care learning: a focus group study, *PLoS ONE* 8 (11) (2013) e80318.
- [3] D.A. Cook, K.J. Sorensen, J.M. Wilkinson, R.A. Berger, Barriers and decisions when answering clinical questions at the point of care: a grounded theory study, *JAMA Intern Med* 173 (21) (2013) 1962–1969.
- [4] G. Del Fiol, T. Workman, P. Gorman, Clinical questions raised by clinicians at the point of care: a systematic review, *JAMA Internal Med* 174 (5) (2014) 710–718.
- [5] M. Dwairy, A.C. Dowell, J.-C. Stahl, The application of foraging theory to information searching behaviour of general practitioners, *BMC Family Practice* 12 (90) (2011).
- [6] D.A. Cook, F. Enders, J.A. Linderbaum, D. Zwart, F.J. Lloyd, Speed and accuracy of a point of care web-based knowledge resource for clinicians: a controlled crossover trial, *Interact J Med Res* 3 (1) (2014) e7.
- [7] J.A. Osheroff, D.E. Forsythe, B.G. Buchanan, R.A. Bankowitz, B. Blumenfeld, R.A. Miller, Physicians' information needs: analysis of questions posed during clinical teaching, *Allans Internal Med* 114 (1991) 576–581.
- [8] Y. Aphinyanaphongs, I. Tsamardinos, A. Statnikov, D. Hardin, C.F. Aliferis, Text categorization models for high-quality article retrieval in internal medicine, *J Am Med Inform Assoc* 12 (2) (2005) 207–216.
- [9] H. Kilicoglu, D. Demner-Fushman, T.C. Rindflesch, N.L. Wilczynski, R.B. Haynes, Toward automatic recognition of high quality clinical evidence, *AMIA Annu Symp Proc* 368 (2008).
- [10] H. Kilicoglu, D. Demner-Fushman, T.C. Rindflesch, N.L. Wilczynski, R.B. Haynes, Towards automatic recognition of scientifically rigorous clinical research evidence, *J Am Med Inform Assoc* 16 (1) (2009) 25–31.
- [11] Y. Aphinyanaphongs, A. Statnikov, C.F. Aliferis, A comparison of citation metrics to machine learning filters for the identification of high quality MEDLINE documents, *J Am Med Inform Assoc* 13 (4) (2006) 446–455.
- [12] E.V. Bernstam, J.R. Herskovic, Y. Aphinyanaphongs, C.F. Aliferis, M.G. Sriram, W.R. Hersh, Using citation data to improve retrieval from MEDLINE, *J Am Med Inform Assoc* 13 (1) (2006) 96–105.
- [13] C. Lokker, R.B. Haynes, N.L. Wilczynski, K.A. McKibbin, S.D. Walter, Retrieval of diagnostic and treatment studies for clinical use through PubMed and PubMed's Clinical Queries filters, *J Am Med Inform Assoc* 18 (5) (2011) 652–659.
- [14] V.M. Montori, N.L. Wilczynski, D. Morgan, R.B. Haynes, T. Hedges, Optimal search strategies for retrieving systematic reviews from Medline: analytical survey, *BMJ* 330 (7482) (2005) 68.
- [15] P. Pirolli, *Information foraging theory: adaptive interaction with information*, Oxford University Press, Oxford, 2007.
- [16] C. Schardt, M.B. Adams, T. Owens, S. Keitz, P. Fontelo, Utilization of the PICO framework to improve searching PubMed for clinical questions, *BMC Med Inform Decis Mak* 7 (2007) 16.
- [17] W.S. Richardson, M.C. Wilson, J. Nishikawa, R.S. Hayward, The well-built clinical question: a key to evidence-based decisions, *ACP J Club* 123 (3) (1995). A12-3.
- [18] A.D. Oxman, D.L. Sackett, G.H. Guyatt, Users' guides to the medical literature. I. How to get started. The evidence-based medicine working group, *JAMA* 270 (17) (1993) 2093–2095.
- [19] A.M. Clark, L.A. Savard, D.R. Thompson, What is the strength of evidence for heart failure disease-management programs?, *J Am Coll Cardiol* 54 (5) (2009) 397–401.
- [20] M. Crandall, T. Duncan, A. Mallat, W. Greene, P. Violano, A.B. Christmas, et al., Prevention of fall-related injuries in the elderly: an eastern association for the surgery of trauma practice management guideline, *J Trauma Acute Care Surgery* 81 (1) (2016) 196–206.
- [21] S.Z. Lewis, R. Diekemper, D.J. Addrizzo-Harris, Methodology for development of guidelines for lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines, *Chest* 143 (suppl 5) (2013) 41s–50s.
- [22] A. Louw, I. Diener, D.S. Butler, E.J. Puenteadura, The effect of neuroscience education on pain, disability, anxiety, and stress in chronic musculoskeletal pain, *Arch Phys Med Rehabil* 92 (12) (2011) 2041–2056.
- [23] J. Ornelas, J.R. Dichter, A.V. Devereaux, N. Kissoon, A. Livinski, M.D. Christian, Methodology: care of the critically ill and injured during pandemics and disasters: CHEST consensus statement, *Chest* 146 (suppl 4) (2014) 35s–41s.
- [24] M.B. Patel, S.S. Humble, D.C. Cullinane, M.A. Day, R.S. Jawa, C.J. Devin, et al., Cervical spine collar clearance in the obtunded adult blunt trauma patient: a systematic review and practice management guideline from the Eastern Association for the Surgery of Trauma, *J Trauma Acute Care Surgery* 78 (2) (2015) 430–441.
- [25] E. Abu-Gharbieh, D.A. Khalidi, M.R. Baig, S.A. Khan, Refining knowledge, attitude and practice of evidence-based medicine (EBM) among pharmacy students for professional challenges, *Saudi Pharm J: SPJ: Off Publ Saudi Pharm Soc* 23 (2) (2015) 162–166.
- [26] R. Brignardello-Petersen, A. Carrasco-Labra, H.A. Booth, M. Glick, G.H. Guyatt, A. Azarpazhooh, et al., A practical approach to evidence-based dentistry: how to search for evidence to inform clinical decisions, *J Am Dent Assoc* (1939) 145 (12) (2014) 1262–1267.
- [27] P.H. Caldwell, T. Bennett, C. Mellis, Easy guide to searching for evidence for the busy clinician, *J Paediatr Child Health* 48 (12) (2012) 1095–1100.
- [28] S. Hosny, M.S. Ghaly, Teaching evidence-based medicine using a problem-oriented approach, *Med. Teach.* 36 (suppl 1) (2014) S62–S68.
- [29] M.J. Rice, Evidence-based practice: a model for clinical application, *J Am Psychiatric Nurses Assoc* 19 (4) (2013) 217–221.
- [30] D. O'Sullivan, S. Wilk, W. Michalowski, K. Farion, Using PICO to align medical evidence with MDs decision making models, *Stud Health Technol Inform* 192 (2013) 1057.
- [31] E.M. LaRue, P. Draus, M.L. Klem, A description of a Web-based educational tool for understanding the PICO framework in evidence-based practice with a citation ranking system, *Comput Inform Nursing: CIN* 27 (1) (2009) 44–49.
- [32] T. Agoristas, A. Merglen, D.S. Courvoisier, C. Combesure, N. Garin, A. Perrier, et al., Sensitivity and predictive value of 15 PubMed search strategies to answer clinical questions rated against full systematic reviews, *J Med Internet Res* 14 (3) (2012) e85.
- [33] Z. Hildon, D. Allwood, N. Black, Impact of format and content of visual display of data on comprehension, choice and preference: a systematic review, *Int J Qual Health Care* (2011) mzzr072.
- [34] P.A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, J.G. Conde, Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support, *J Biomed Inform* 42 (2) (2009) 377–381.
- [35] J. Brooke, SUS – a quick and dirty usability scale, *Usabil Eval Indust* 189 (184) (1996) 4–7.
- [36] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, *J Roy Stat Soc: Ser B (Methodol)* (1995) 289–300.
- [37] S.Z. Shariff, J.M. Sontrop, R.B. Haynes, A.V. Iansavichus, K.A. McKibbin, N.L. Wilczynski, et al., Impact of PubMed search filters on the retrieval of evidence by physicians, *CMAJ* 184 (3) (2012) E184–E190.
- [38] R. Mishra, J. Bian, M. Fiszman, C.R. Weir, S. Jonnalagadda, J. Mostafa, et al., Text summarization in the biomedical domain: a systematic review of recent research, *J Biomed Inform* 52 (2014) 457–467.
- [39] S.E. Rosenbaum, C. Glenton, H.K. Nyland, A.D. Oxman, User testing and stakeholder feedback contributed to the development of understandable and useful Summary of Findings tables for Cochrane reviews, *J Clin Epidemiol* 63 (6) (2010) 607–619.
- [40] F.D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS Quart* 13 (3) (1989) 319–340.
- [41] J.W. Ely, J.A. Osheroff, M.L. Chambliss, M.H. Ebell, M.E. Rosenbaum, Answering physicians' clinical questions: obstacles and potential solutions, *J Am Med Inform Assoc* 12 (2) (2005) 217–224.
- [42] W. Hsu, W. Speier, R.K. Taira, Automated extraction of reported statistical analyses: towards a logical representation of clinical trial literature, *AMIA Annu Symp Proc* 2012 (2012) 350–359.
- [43] Summerscales RL, Argamon S, Bai S, Huperff J, Schwartzff A, editors. Automatic summarization of results from clinical trials. In: 2011 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2011; 12–15 Nov. 2011.