Usability Methods and Evaluation Criteria for Published Clinical Guidelines on the Web: A Systematic Literature Review

Soudabeh Khodambashi[(✉)] and Øystein Nytrø

Norwegian University of Science and Technology, Trondheim, Norway
{Soudabeh, Nytroe}@ntnu.no

Abstract. The usability evaluation of published clinical guidelines (GL) on the web is an important analytical tool. This evaluation helps to determine how presentation affects GL use; it identifies the user's needs and assesses whether the user's perceived success rate in finding an answer is reliable or not. Such information is of great value since an inaccurate perceived success rate could lead to potentially critical consequences. This paper explores literature focusing on the usability evaluation of GL web-sites. We examine the evaluation goal, criteria and methods that researchers considered in GL website evaluation. We found that although many researchers have concentrated on the evaluation of clinical decision support systems and their usability; a problem subsists. Evaluation of the usability of published GLs on the Web and the understanding of the users' interaction is in its infancy. Building GL websites is not substantially different than building a highly functional website with high usability in general. However, there are particular factors such as clinicians' time constraints and information overload that need to be considered in the design of a GL website and its evaluation.

Keywords: Clinical guidelines · Guideline website · Evaluation · Usability

1 Introduction

Clinical guidelines (GLs) are widely used. According to the National Guidelines Clearinghouse, more than 320 organizations have been involved in the development of GLs by 2016, including CGLs authoring organizations, academic research groups, and commercial publishers. There is no real length restriction for GLs and they are produced in a variety of digital and print formats. They can be as short as a single page or as long as a booklet of more than 15 pages. Presenting the GL content in PDF format reduces the chance of finding a relevant answer quickly [1]. To increase accessibility, some GL authoring organizations publish their documents on the Web.

Studies on the accessibility of GLs have revealed that clinicians need to find answers to their questions within 2 min [2]. It is therefore highly important that relevant, accurate information is available as quickly and easily as possible and well within this time limit. However, publishing GLs on the Web may not necessarily make them easier to use, as a poor design can make a system difficult to learn and complicated to use and leads to negative consequences [3]. Furthermore finding the right information

in a GL and comparing GLs of interest can be challenging for users, especially as the number of electronically available GLs increases. Therefore, a usability evaluation of published GLs on the Web is necessary to investigate how presentation affects GL use. In this paper we systematically reviewed the literature to summarize the existing publications on usability evaluation of published GLs on the Web considering evaluation criteria, metrics and evaluation methods.

2   Materials and Methods

To retrieve and extract data from relevant studies, we performed a systematic literature review using PubMed and Google Scholar databases. The selection process is presented in Fig. 1. The last search was conducted in March 2017. In order to find more relevant literature, we used a backward snowballing method. Note that our literature search did not include clinical decision support system and their usability evaluation. We employed a Thematic analysis method [4] to identify the evaluation themes presented in Table 1: (1) Usability, (2) Using iconic language, (3) Searching, and (4) Patients guidelines.
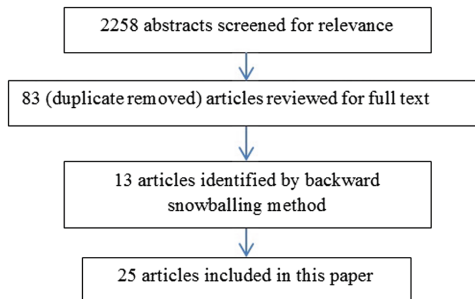

Fig. 1.   Selection process of retrieved articles

3   Results

Table 1.   Identified articles, their evaluation criteria and methods

Ref.	Evaluation goal/criteria	Method
Theme 1: Usability		
[5]	Evaluation of GL presentation in XML, PDF, and PDA	Evaluation of the satisfaction of GL general users, GL developers and GL reviewers by asking questions
[6]	Evaluating website quality to identify factors affecting health care workers' adoption of GL website: accuracy, completeness, readability,	Semi-structured interviews, the inter-views were based on the PRECEDE (predisposing,

<div align="right">(continued)</div>

Table 1. (*continued*)

Ref.	Evaluation goal/criteria	Method
	design, provided references, disclosures, usability, findability, relevance	reinforcing, and enabling causes in educational diagnosis and evaluation) model
[7]	User experiences of GLs on mobile through the concept of webflow: navigation, learning, focused attention, challenges, orientation	Data from users (who installed the mobile app) was collected through the online questionnaire
[8]	How do GL features influence their use?	Literature search
[9]	Evaluation of acceptability and usefulness	Initial online survey and a more de-tailed follow-up feedback survey emailed to web users
[10]	To assess the effect of differing GL representation formats on the quality of nursing care plans and on the experiences of nurses	Scenario-based and task completion. The GLs were presented in two for-mats: PDF and web based interactive. Participants were asked to 'think-aloud' during task completion and their experiences were recorded, transcribed, and analysed through a cognitive task analysis
[11]	To fine tune the presentation of GL information	Remote collection of both quantitative logging data (browsing) and qualitative use (on user preferences, information) and usability issues from users of GL system
[12]	Case studies on website look and feel	Interviews
[13]	Evaluate how the structure of GLs accompanied with search function impacted finding the right answer, GL usage and efficiency: response accuracy, users' satisfaction and performance	Presenting GLs to the participants (divided in two groups) in two different ways. Survey (questionnaire), scenario based task completion, questionnaire to collect feedback on the GL structure, ease of finding the answer, advantages of the webpage, and their experience with search functionality
[14, 15]	Testing multi-layered presentation format of GLs on their developed prototype: findability, usefulness, usability, understandability, credibility, and desirability	User testing, semi structured interviews (on the overall structure, layout, and components of the format), applying a think-aloud method for exploring important aspects of user experience
[16]	Users' interaction and performance: efficiency, effectiveness, learnability, response accuracy, number of mouse clicks and usage rate for search functions, task completion time, users' objective and perception of task success rate, and learning effect for inexperienced users	Five GL websites were evaluated and compared using an eye-tracker, a preliminary survey, a scenario-based task completion, and a semi-structured interview
[17]	Usability evaluation of five GL websites and users' feedback. Metric: perceived usability	A pretest survey, scenario-based task completion, system usability scale (SUS) questionnaire, observation, and semi-structured interview
[18]	To evaluate the efficacy, acceptability and feasibility of using QR codes to facilitate 'Just in Time' learning of GLs by measuring usage statistics such as page views, unique page views and average time spent on page	Website analytics and semi-structured interviews
[19]	Comparing different methods of GL dissemination: "health professionals' perceived" usability and practice behaviour change of information and communication technologies	Systematic literature survey

Table 1. (*continued*)

Ref.	Evaluation goal/criteria	Method
[14, 20]	To investigate physicians' preferences, perceived usefulness and understanding of a new multi-layered GL presentation format compared to a standard format	View random clinical scenario and GL recommendation in a multi-layered format or standard format to physicians after which they answered multiple-choice questions using clickers
Theme 2: Using iconic language		
[21]	To assess VCM, if the language is easy to learn, understand and use. Respondents' document length and question type were documented and evaluated	Participants were asked to register VCM training time, to indicate the meaning of VCM icons and sentences, and to answer clinical questions related to randomly generated drug monograph-like documents, supplied in text or VCM format. Compared the correctness of responses and the response times obtained with text and VCM and applied linear regression analysis
[22]	Usability study of an iconic user interface to ease information retrieval of GLs, comparing a Visualization of Concepts (VMC) with a non-VCM inter-face: time taken, users' ability, and perceived usefulness	Scenario-based (two different scenarios for each interface). The ability and time taken to select a relevant re-source were recorded and compared. A usability analysis was performed using SUS
[23]	To evaluate VCM for the consultation of GLs: response times, number of errors, response accuracy, perceived usability	Comparison of response times, response accuracy and the number of recorded errors during task completion using VCM or a textual interface. Users' perceived usability was evaluated with SUS
Theme 3: Searching		
[24]	Comparing concept-based and context-sensitive GL search in free-text search retrieval performance	Precision and recall of the designed search engines
[25]	Health information-seeking behaviour on the Web: internet use and ascertaining challenges	Literature review on the topic area from 2006 to 2010
[26]	Comparing user experiences and perceived usability on two proto-types: search-based and content-based recommendation ranking of GLs	A survey (pretest questionnaire), scenario-based with given tasks, SUS and interview
[27]	Information searching behaviour of medical students, evaluated the effect of varying levels of task difficulty on search behaviour according to demographic variables. Querying details, search results interaction details, querying versus clicking behaviour and task completion time were evaluated	Participants were attended in an inter-active information retrieval experiment type methodology that was used to study the interactive searching behaviour with structured observation
Theme 4: Patients guidelines		
[28]	To assess their Portal's functionality, effectiveness and identify any usability problems from perspective of the patients: the quality of the provided information, whether the information they accessed had helped in any decisions they had to make, and the preferred search options	Two focus groups reviewed the usefulness of the Portal, 6 women participated in the pilot usability evaluation, and 13 women participated in the onsite usability evaluation
[29]	User test of a patient version of a SIGN GLs	Using a think-aloud protocol method, all sessions were recorded and transcribed

4 Discussion and Conclusion

According to the results, no gold standard has been considered in the GL website evaluations. We identified that measuring efficiency was the most used criteria in GL evaluation including task completion time, time spent and number of made errors. The second most popular evaluative criteria was perceived usefulness by applying SUS method, followed by presentation format. The number of evaluations focusing on usability and usefulness, however reviewing the articles revealed that it is not clear how they evaluated them. It is necessary for researchers clearly report how they evaluate and measure usefulness and usability. Although searching function is one of the important factors in findability of information on a GL website, not much attention has been paid to it. GL websites should not only be assessed by ease of use, presentation format, layout, and supported digital features with intuitive and simplified navigation, but also it is necessary that efficient search and the format of search results presentation are evaluated. As clinicians' time constraints and information overload are two factors in GLs adoption, evaluation of the search function and its retrieval performance in efficiently identifying relevant GLs is needed (i.e. a trained search function for clinical terms, especially for synonyms, acronyms, and abbreviations).

References

1. Green, M.L., Ciampi, M.A., Ellis, P.J.: Residents' medical information needs in clinic: are they being met? Am. J. Med. **109**, 218–223 (2000)
2. Coumou, H.C., Meijman, F.J.: How do primary care physicians seek answers to clinical questions? A literature review. J. Med. Libr. Assoc. **94**, 55 (2006)
3. Jaspers, M.W.M.: A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. Int. J. Med. Inform. **78**, 340–353 (2009)
4. Cruzes, D.S., Dyba, T.: Recommended steps for thematic synthesis in software engineering. In: 2011 International Symposium on Empirical Software Engineering and Measurement (ESEM), pp. 275–284 (2011)
5. Park, M.: Development and evaluation of online evidence based guideline bank system. Stud. Health Technol. Inform. **122**, 105 (2006)
6. Verhoeven, F., Steehouder, M.F., Hendrix, R.M., van Gemert-Pijnen, J.E.: Factors affecting health care workers' adoption of a website with infection control guidelines. Int. J. Med. Inform. **78**, 663–678 (2009)
7. Oinas-Kukkonen, H., Raisanen, T., Leiviska, K., Seppanen, M., Kallio, M.: Physicians' user experiences of mobile pharmacopoeias and evidence-based medical guidelines. Int. J. Healthc. Inf. Syst. Inform. (IJHISI) **4**, 57–68 (2009)
8. Gagliardi, A.R., Brouwers, M.C., Palda, V.A., Lemieux-Charles, L., Grimshaw, J.M.: How can we improve guideline use? A conceptual framework of implementability. Implement. Sci. **6**, 26 (2011)
9. Berk, L., Berk, M., Dodd, S., Kelly, C., Cvetkovski, S., Jorm, A.F.: Evaluation of the acceptability and usefulness of an information website for caregivers of people with bipolar disorder. BMC Med. **11**, 162 (2013)

10. Csima, D.G.: The effect of clinical practice guideline representation on nursing care planning, University of Victoria, Canada (2013)
11. Kushniruk, A., Kaipio, J., Nieminen, M., Hyppönen, H., Lääveri, T., Nohr, C., Kanstrup, A.M., Christiansen, M.B., Kuo, M.-H., Borycki, E.: Human factors in the large: experiences from Denmark, Finland and Canada in moving towards regional and national evaluations of health information system usability: contribution of the IMIA human factors working group. Yearb. Med. Inform. **9**, 67 (2014)
12. Horvath, K.J., Ecklund, A.M., Hunt, S.L., Nelson, T.F., Toomey, T.L.: Developing internet-based health interventions: a guide for public health researchers and practitioners. J. Med. Internet Res. **17**, e28 (2015)
13. Khodambashi, S., Wang, Z., Nytrø, Ø.: Reality versus user's perception in finding answer to clinical questions in published national guidelines on the web: an empirical study. Proced. Comput. Sci. **63**, 268–275 (2015)
14. Kristiansen, A.: Dissemination and adaptation strategies customized for trustworthy practice guidelines using the GRADE framework (2016)
15. Kristiansen, A., Brandt, L., Alonso-Coello, P., Agoritsas, T., Akl, E.A., Conboy, T., Elbarbary, M., Ferwana, M., Medani, W., Murad, M.H.: Development of a novel, multilayered presentation format for clinical practice guidelines. CHEST J. **147**, 754–763 (2015)
16. Khodambashi, S., Gilstad, H., Nytrø, Ø.: Usability evaluation of clinical guidelines on the web using eye-tracker. In: Medical Informatics Europe (MIE 2016) (2016)
17. Khodambashi, S., Nytrø, Ø.: Usability evaluation of published clinical guidelines on the web: a case study. Accepted-under publication (2016)
18. Jamu, J.T., Lowi-Jones, H., Mitchell, C.: Just in time? Using QR codes for multi-professional learning in clinical practice. Nurse Educ. Pract. **19**, 107–112 (2016)
19. De Angelis, G., Davies, B., King, J., McEwan, J., Cavallo, S., Loew, L., Wells, G.A., Brosseau, L.: Information and communication technologies for the dissemination of clinical practice guidelines to health professionals: a systematic review. JMIR Med. Educ. **2** (2016)
20. Brandt, L., Vandvik, P.O., Alonso-Coello, P., Akl, E.A., Thornton, J., Rigau, D., Adams, K., O'Connor, P., Guyatt, G., Kristiansen, A.: Multilayered and digitally structured presentation formats of trustworthy recommendations: a combined survey and randomised trial. BMJ Open **7**, e011569 (2017)
21. Lamy, J.-B., Duclos, C., Bar-Hen, A., Ouvrard, P., Venot, A.: An iconic language for the graphical representation of medical concepts. BMC Med. Inform. Decis. Mak. **8**, 16 (2008)
22. Griffon, N., Kerdelhué, G., Hamek, S., Hassler, S., Boog, C., Lamy, J.-B., Duclos, C., Venot, A., Darmoni, S.J.: Design and usability study of an iconic user interface to ease information retrieval of medical guidelines. J. Am. Med. Inform. Assoc. **21**, e270–e277 (2014)
23. Pereira, S., Hassler, S., Hamek, S., Boog, C., Leroy, N., Beuscart-Zéphir, M.-C., Favre, M., Venot, A., Duclos, C., Lamy, J.-B.: Improving access to clinical practice guidelines with an interactive graphical interface using an iconic language. BMC Med. Inform. Decis. Mak. **14**, 77 (2014)
24. Moskovitch, R., Shahar, Y.: Vaidurya: a multiple-ontology, concept-based, context-sensitive clinical-guideline search engine. J. Biomed. Inform. **42**, 11–21 (2009)
25. Barry, M.M., Domegan, C., Higgins, O., Sixsmith, J.: A literature review on health information seeking behaviour on the web: a health consumer and health professional perspective (2011)
26. Khodambashi, S., Perry, A., Nytrø, Ø.: Comparing user experiences on the search-based and content-based recommendation ranking on stroke clinical guidelines-a case study. Proced. Comput. Sci. **63**, 260–267 (2015)

27. Inthiran, A., Alhashmi, S.M., Ahmed, P.K.: A user study on the information search behaviour of medical students. Malays. J. Libr. Inf. Sci. **20**, 61–77 (2015)
28. McKemmish, S., Manaszewicz, R., Burstein, F., Fisher, J.: Consumer empowerment through metadata-based information quality reporting: the breast cancer knowledge online portal. J. Assoc. Inf. Sci. Technol. **60**, 1792–1807 (2009)
29. Fearns, N., Graham, K., Johnston, G.: Improving the user experience of patient versions of clinical guidelines: user testing of a Scottish Intercollegiate Guideline Network (SIGN) patient version. BMC Health Serv. Res. **16**, 37 (2016)