

RESEARCH

Open Access



Evaluating the usability of a cancer registry system using Cognitive Walkthrough, and assessing user agreement with its problems

Fatemeh Bagheri¹, Faezeh Abbasi¹, Mojtaba Sadeghi² and Reza Khajouei^{1*}

Abstract

Objective/aim Good design of cancer registry systems makes them easy to use, while poor design of their user interfaces leads to user dissatisfaction and resistance. The objective of this study was to evaluate the usability of a cancer registry system using Cognitive Walkthrough (CW) and to assess users' agreement with its usability problems.

Methods CW was used to evaluate the registry system. We developed a checklist to help evaluators speed up the evaluation process, a problems form to collect the usability issues identified by the evaluators, and a problems severity form to determine the severity of problems by the evaluators. The problems were classified into two categories according to the CW questions and the system tasks. The agreement of the users with the system problems was examined by an online questionnaire. Users' agreement with the problems was then analyzed using the Interclass Correlation Coefficient in the SPSS 22 (Statistical Package for Social Science).

Results In this study, 114 problems were identified. In the categorization of problems based on the CW questions, 41% (n = 47) of the problems concerned the issue of "users do not know what to do at each stage of working with the system", 24% (n = 27) were classified as "users cannot link what they intend to do with system controls", and 22% (n = 25) were related to "user's lack of understanding of the system processes". Based on user tasks, about 36% (n = 41) of the problems were related to "removing patient duplication" and 33% (n = 38) were related to "registration of patient identification information". User agreement with the problems was high (CI 95% = 0.9 (0.96, 0.98)).

Conclusion System problems often originate from user ignorance about what to do at each stage of using the system. Also, half of the system problems concern a mismatch between what users want to do and the system controls, or a lack of understanding about what the system does at different stages. Therefore, to avoid user confusion, designers should use clues and guides on the screen for users, design controls consistent with the user model of thinking, and provide appropriate feedback after each user action to help users understand what the system is doing. The high agreement of users with the problems showed that in the absence of users system designers can use CW to identify the problems that users face in the real environment.

Keywords Usability, Cognitive walkthrough, Registry system, User experience

*Correspondence:

Reza Khajouei
r.khajouei@kmu.ac.ir; r.khajouei@yahoo.com

¹ Department of Health Information Sciences, Faculty of Management and Medical Information Sciences, Kerman University of Medical Sciences, Kerman, Iran

² Department of Health Information Technology, Faculty of Paramedicine, AJA University of Medical Sciences, Tehran, Iran



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

According to the World Health Organization, cancer is the second leading cause of death after cardiovascular disease worldwide [1]. The first step in controlling cancer is gathering information about the incidence, type, and location of the disease [2]. In this regard, the cancer registration system was developed. This system helps to better manage and control the disease by providing accurate information about its prevalence and patient survival [3].

One of the factors contributing to the success of an information system is its good design [4]. Good design is consistent with the goals, needs, and interests of users [5]. Given the difference in computer experience and expertise of users and the unfamiliarity of information systems developers with the user population, the design of these systems is not often tailored to the needs and interests of their users. Also, according to previous studies [6–9], poor design of information systems leads to users' fatigue, and dissatisfaction, as well as, system inefficiency and rejection. The good design of an information system contributes to its usability. According to ISO 9241-11, usability means that a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [10]. Usability refers to aesthetic appearance, ease of use, easy navigation, user-friendliness, and clarity of the design of an information system [11]. Despite the increasing focus on usability and the development of relevant methods and procedures, information systems designers fail to design systems based on usability principles [9].

Usability evaluation methods fall into two user-based and expert-based categories. In the user-based evaluation methods, users participate in the evaluations to identify problems with information systems. Since recruiting a sufficient number of real users for user-based usability evaluation methods is sometimes difficult or expensive, it is essential to use other methods that identify usability problems hindering real users [12, 13]. In expert-based evaluation methods, experts identify usability problems in view of users' abilities [14]. Cognitive Walkthrough (CW) is one of the most common expert-based evaluation methods [15]. The main purpose of CW is to identify the simplest way to perform a task by creating an action sequence tree for the task and its sub-tasks. In this method, evaluators identify most usability problems of an interface and their causes especially severe problems easily, quickly, and at a low cost. Also, this method is suitable for evaluating information systems that their users are beginners [16]. However, it is unclear how much information system users agree with the problems identified by evaluators in such expert-based methods. Thus, examining the agreement of users with usability problems

helps to improve the accuracy of the results of the expert-based evaluation methods, especially the CW method. So far, the agreement of information system users with the problems identified in expert-based usability evaluations has not been sufficiently studied. Among previous studies, only one study [17] measured the extent to which users agreed with the problems identified by the evaluators in a Heuristic Evaluation of an emergency information system. Other similar studies reported the results of CW evaluations [18–20], developed CW extensions such as Heuristic Walkthrough [21], compared CW with other expert-based methods such as Heuristic Evaluation [6, 22], compared CW with user-based methods [23, 24], or compared other expert-based evaluation methods, including Heuristic Evaluation with user-based methods [25, 26]. One of the information systems in the healthcare domain is the cancer registry system, which has been developed to improve the implementation and management of cancer registration programs. Good design and high usability of this system result in better effectiveness, efficiency, and user satisfaction of the system for users [10]. Therefore, the objective of this study was to evaluate the usability of a cancer registry system using CW and to assess users' agreement with the problems identified by CW.

In this study, first, the usability problems of a cancer registry system were identified using CW. Next, the severity of the problems was determined by the evaluators using a severity rating scale. Then, the problems with higher severities were given to users for determining the severities by this group. Finally, users' agreement with the problems identified by CW was assessed.

Research methodology

System and setting

The cancer registry system is a national information system that was implemented in 2015 and is currently used by more than one hundred users in twenty universities in Iran. The cancer registry system allows the collection and management of information on cancer patients and the calculation of the incidence and prevalence of cancer in different geographical areas. Registration of patients' profiles and their tumor information, authentication, removing patient duplications, and annual reporting of patients' information are among the capabilities of this system. The most frequent and routine tasks of the system are "Recording patient's personal details and tumor information" and "Removing patient duplications". The present study was performed in the cancer registration center of Kerman University of Medical Sciences (KUMS).

This study was conducted in two phases:

Phase 1. Evaluation of the cancer registry system using CW

The CW method introduced by Polson and Lewis [27, 28] was used for the evaluation. This phase consisted of three steps: 1. evaluation preparation, 2. evaluation execution and 3. analyzing the problems. In the preparation step, with the assistance of an expert familiar with the workflow of the cancer registry system, two scenarios were developed based on two frequent and routine tasks of the system users (Fig. 1). Performing these scenarios in the system required the completion of six sub-tasks. The sub-tasks and the actions required to complete each sub-task in the system are listed in Fig. 2. Finally, to speed up the evaluation process and to reduce the likelihood of error, a checklist was developed based on the CW method for evaluating the execution of the sub-tasks and their actions in the system. Then, to confirm the validity of prepared scenarios and defined sub-tasks, this checklist was reviewed by an expert familiar with the cancer registry system and a medical informatics expert. This checklist included the sub-tasks, the user’s purpose for completing the sub-tasks, the actions needed to complete the sub-tasks, and the four questions recommended to be

asked for the evaluation of each action in CW (1. Will the user try to achieve the right effect?, 2. Will the user notice that the correct action is available?, 3. Will the user associate the correct action with the effect that the user is trying to achieve?, and 4. If the correct action is performed, will the user see that progress is being made toward the solution of the task?) [20]. Overall, this checklist included 6 sub-tasks, 139 actions, and 556 questions (139 * 4) (Fig. 3) (see Additional file 1).

In the second step, since a sample size of 10 ± 2 is sufficient to detect 80% of the usability problems of a system [30–33], eight evaluators were invited to perform the evaluation. These evaluators had a background in medical informatics and health information technology and received at least six months of training and experience in using CW. Before the evaluation, evaluators held a joint meeting with an evaluation expert having more than ten years of experience in evaluating health information systems. In this meeting, the procedure of evaluation was discussed and all evaluators worked with the cancer registry system and got acquainted with the routine process of executing scenarios in the system. To begin the evaluation, paper-based copies of

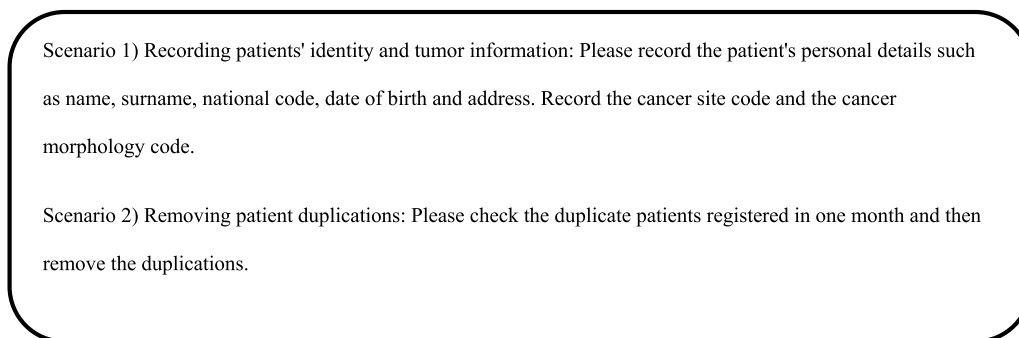


Fig. 1 Two routine scenarios in cancer registration

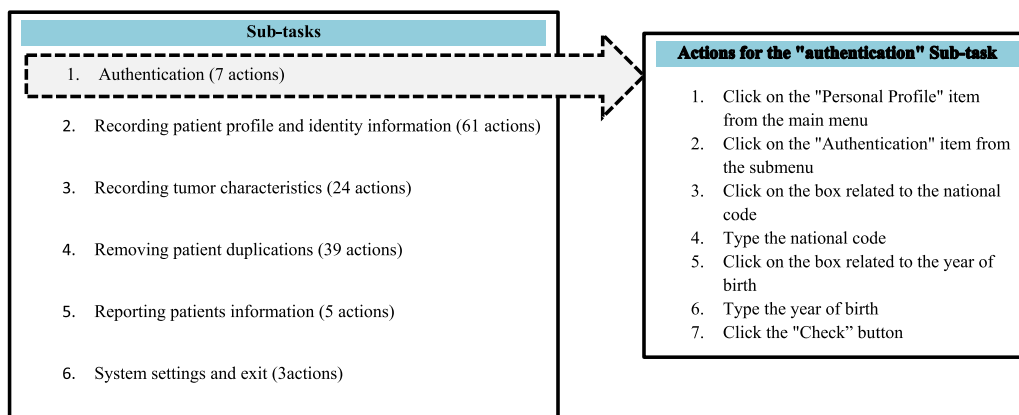


Fig. 2 List of sub-tasks related to the two scenarios and an example of actions requires for completing the "Authentication" sub-task

Sub task 1) Please authenticate the patient.

Goals	actions	Questions			
		1. Will the user try to achieve the right effect?	2. Will the user notice that the correct action is available?	3. Will the user associate the correct action with the effect that user is trying to achieve?	4. If the correct action is performed, will the user see that progress is being made toward solution of the task?
a) Login to the authentication page	1) Click on the "Personal Profile" item from the main menu				
	2) Click on the "Authentication" item from the submenu				
	3) Click on the box related to the national code				
	4) Type the national code				
	5) Click on the box related to the year of birth				
	6) Type the year of				

Fig. 3 A part of the checklist developed to help evaluators speed up the evaluation process and to reduce the likelihood of error

the checklist were distributed among the evaluators. The evaluators independently performed the sub-tasks listed in the checklist, took the actions one by one, and determined how easily users can work with the system, taking the system's design and feedback, as well as, the users' background and knowledge into account. The answer to the checklist questions was "yes" if no problem was detected, and "no" if there was a problem in the system design. Whenever the answer to any of the questions was no, the evaluators entered a description of the problem into the checklist. In the third step, a problems form was designed to collect the problems identified by the evaluators. In this form, each problem was described with six features including sub-task, action, question number, evaluator code, error location, and error description (Fig. 4) (see Additional file 2). To analyze the problems, the questions with yes answers by all evaluators were discarded, but questions to which at least one of the evaluators gave a no answer were included in the problems form for further consideration. Finally, by combining duplicate items in the problems form, the final problem list was formed.

To determine the severity of usability problems of information systems, among several severity rating method [34–36], we used a common severity rating method proposed by Nielsen that prioritizes the problems in terms of their effect on user interaction [31, 36–39]. In this method, the severity of problems is determined based on three criteria of frequency, impact, and persistence [40]. Thus, to determine the severity of problems by the

Sub-task	Action	Question number	Evaluator code	Problem location	Problem description
Authentication	1) Click on the "Personal Profile" item from the main menu	1	1, 6	Main menu	The user does not know the place of authentication in the system.
Authentication	2) Click on the "Authentication" item from the sub menu	4	1, 2, 3, 6	Authentication page	The required feedback was not received.

Fig. 4 A part of the problems form used for collecting problems identified by the evaluators

evaluators, a problems severity form was developed by the researchers based on Nielsen severity rating. This form consisted of two parts. The first part was a guide for determining the severity of the problems according to Nielsen severity rating. The second part included the problem number, problem location, problem description, question number, number of evaluators who identified the problem, problem severity, and solution. Using this form, evaluators independently assigned a score from 0 to 4 to each problem (0=not a problem. 1=cosmetic problem, does not need to be fixed. 2=minor problem, low priority for fixing 3=major problem, high priority for fixing. 4=usability catastrophe, imperative to be fixed) (Fig. 5) (see Additional file 3). To determine the final severity of each problem, the mean severity of that problem was calculated. In the end, the problems were categorized based on evaluation questions. For example for the question, “Will the user try to achieve the right effect?” the category was created under the heading “Notifying of action at hand”. Therefore, the problems were divided into four categories: “Notifying of action at hand”, “Existence of necessary controls”, “Visibility and clarity of controls for users” and “User perception of the work process in the system”. To be able to prioritize the tasks based on their difficulty and their importance for fixing, the problems were also categorized by tasks. Upon

this categorization, the problems were divided into five categories: “Authentication”, “Recording patient identity information”, “Recording tumor characteristics”, “Removing patient duplications” and “Reporting patient information”. To investigate how the complexity of a sub-task can hinder the successful completion of that sub-task by users, the relationship between the number of actions in each sub-task and the number of problems users encountered when performing the sub-task was tested using Spearman’s test.

Phase 2. Examining user agreement with the problems identified by the evaluators in the cancer registry system

An online questionnaire was developed (in Google Forms) to assess the agreement of users with problems identified in this study. Due to the large number of problems, and to increase the response rate of users the severity cut-off was set at 3.5 so that only problems with higher severities (severity > 3.5) were included in the questionnaire (n = 32). This questionnaire had two sections. The first section included ten questions concerning age, gender, education, computer-related degree, ability to work with a computer, experience of using personal computers, the rate of internet use, work experience in cancer registration, and the experience of working with cancer registry system and other health

Section 1) Dear evaluator please read the following guide for determining the severity of problems

The severity of a usability problem is a combination of three factors:

- The **frequency** with which the problem occurs: Is it common or rare?
- The **impact** of the problem if it occurs: Will it be easy or difficult for the users to overcome?
- The **persistence** of the problem: Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?

Use the following 0 to 4 rating scale to rate the severity of usability problems:

0 = I don't agree that this is a usability problem at all

1 = Cosmetic problem only: need not be fixed unless extra time is available on a project

2 = Minor usability problem: fixing this should be given low priority

3 = Major usability problem: important to fix, so should be given high priority

4 = Usability catastrophe: imperative to fix this before a product can be released

Section 2) Dear evaluator please determines the severity and a solution for each identified problem.

problem number	problem location	problem description	question number	number of evaluators who identified the problem	problem severity	proposed solution

Fig. 5 A part of the problems severity form

information systems. Except for the age, the answers to the rest of the questions were multiple choice. The second section contained 32 problems with high severity. These problems were multiple choice with zero (no problem) to four (catastroph problem, imperative to fix) options. To better present and clarify the location of each problem, a screenshot of the system where the problem occurred was placed after each problem. The content and face validity of the questionnaire were confirmed by three medical informatics specialists. The reliability of the questionnaire was confirmed by Cronbach's alpha of 93% (Fig. 6) (see Additional file 4). The purpose of usability studies is to identify the problems that users encounter during interaction with an information system. CW is a task-based method that focuses on the usability of system for users who would work directly with the system. Since, IT experts contribute to the development and maintenance of information

systems and health professionals are not familiar with most of the tasks of information systems, this study has only focused on the real users of the system. Thus, the participants in this stage were users of the cancer registry system who work in the cancer registration centers. The following two criteria were used to select the study participants, (1) having more than a month of work with the cancer registry system, and (2) voluntary participation. To collect the data, the participating users were provided with the questionnaire link via WhatsApp or SMS (Short Message Service). The purpose of the study was mentioned in the introduction of the questionnaire, and the participants were ensured that the confidentiality of their information would be maintained. To obtain informed consents, the participants were asked to answer the questionnaire only if they were willing to participate in the study (Fig. 7).

The user does not know how to enter the code of the type, and location of the disease ?

0 1 2 3 4

No problem catastrophic problem

مرکز Icdو

مکان Icdو

Stage M Stage N Stage T

Not applicable Not applicable Not applicable

مرکز Icdو

Clinical Stage

تکرار آسان

پرسش خودروبرای مشاهده

Fig. 6 A sample of the questions used to assess the agreement rate of users with problems identified by CW

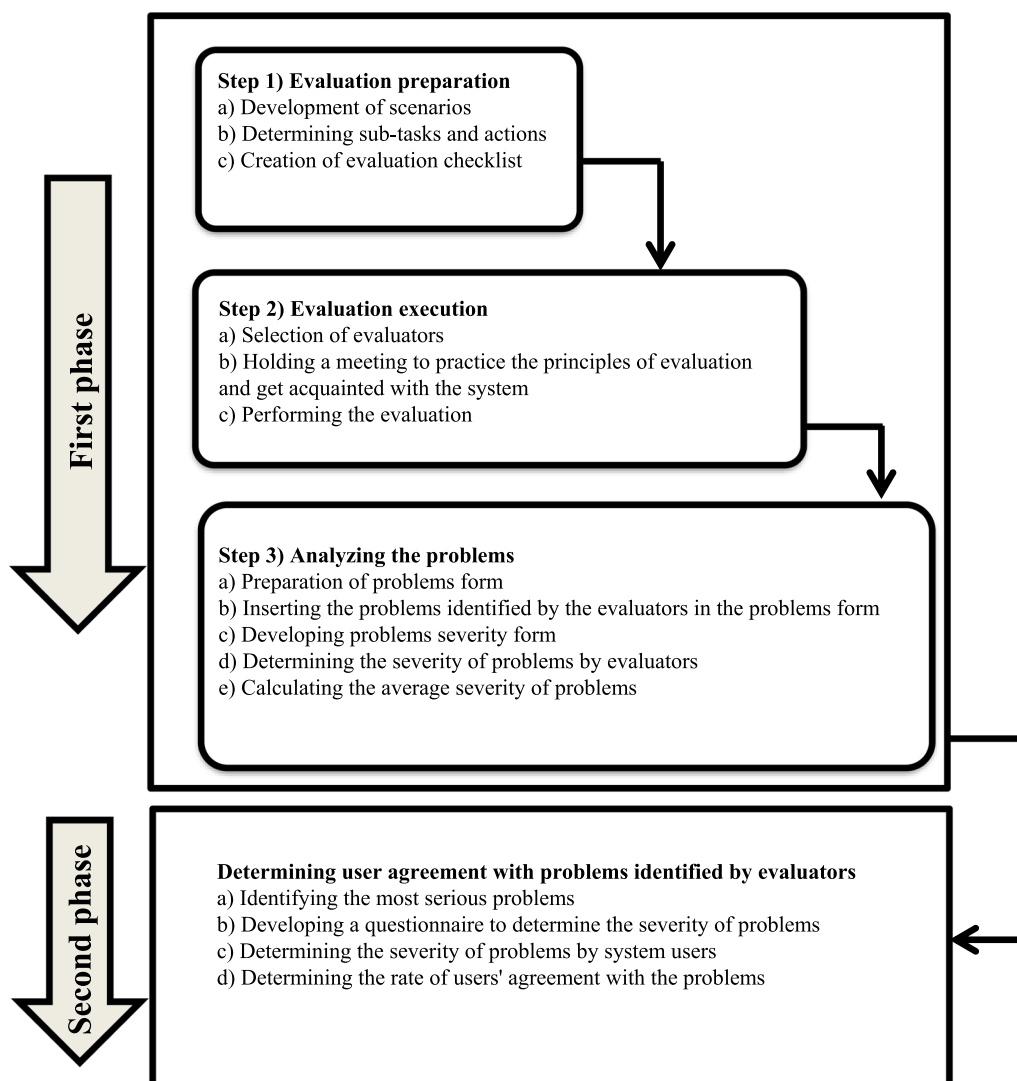


Fig. 7 The study workflow of CW evaluation and assessing user agreement with the identified problems

To analyze the data for this step, the data were extracted from Google Forms in Excel format and imported into SPSS 22 (Statistical Package for Social Science). The Interclass Correlation Coefficient was then used to calculate the coefficient of participant agreement using the mean scores that participants gave to each question and the mean severities of these issues. The study workflow is shown in Fig. 7.

Results

Phase 1. The CW evaluation of the cancer registry system

The evaluators’ demographic information is shown in Table 1. Most of the evaluators (75%) were women, less than 30 years old, and had a master’s degree. Also, more than half (62%) of the evaluators were familiar with

health information systems and had the experience with information systems evaluation.

Evaluators identified 435 problems during CW of the registry system. After merging problems and removing duplicates, 114 problems remained (see Additional file 5).

Categorizing problems based on evaluation questions

The categorization of problems based on evaluation questions along with the number and mean severity of problems are shown in Table 2. About 90% of the problems were minor or major. Most of the problems were related to the “notifying of action at hand” category. The categories “Visibility and clarity of controls for users” and “User perception of the work process in the system” each accounted for approximately 25% of the problems.

Table 1 Evaluators’ demographic information

Demographic information	n (%)
<i>Age</i>	
< 30	6 (75)
>= 30	2 (25)
<i>Sex</i>	
Male	2 (25)
Female	6 (75)
<i>Education degree</i>	
Master’s	6 (75)
PhD	2 (25)
<i>Acquaintance with health information systems</i>	
Medium	3 (37.50)
High	5 (62.50)
<i>Experience of information systems evaluation</i>	
Medium	3 (37.50)
High	5 (62.50)

The mean severity of problems in these categories was almost at the same level, between 1.9 and 2.2. The lowest number of problems and also the lowest mean severity of problems were related to the “existence of necessary controls” category.

Categorizing problems based on the scenarios

The categorization of the problems by sub-tasks, along with the number and mean severity of problems, is shown in Table 3. The mean severity of the problems in each sub-task was at almost the same level, between 1.8 and 2.6. The highest number of problems were related to the “Removing patient duplications” and “Recording patient identity information” sub-tasks. Using the Spearman test, no correlation was observed between the number of actions and the number of problems per task ($p > 0.05$, Spearman correlation = 0.75).

Phase 2. Examining the agreement of users with the problems identified by CW

Out of 101 cancer registry users, 27 (26.73%) users responded to the questionnaire. The demographic information of these twenty-seven participants is shown in Table 4. In general, about 70.4% (n=19) of the participants were female and a third of the participants were younger than 30 and a third were between 40 and 49 years old. More than half of the participants (51.8%) had a bachelor’s degree, computer-related education (55.6%), and an advanced level of computer skills (62.9%). About 37% of the participants (n=10) stated that they use a personal computer daily and about 85% (n=23) stated that they use the Internet for two to several hours a day. Forty-four percent of the participants (n=12) had between 2 and 10 years of experience in the cancer

Table 2 Categorization of problems based on evaluation questions and mean problem severity

Category based on evaluation questions	n (%)	Severity n (%)				Average severity
		Cosmetic	Minor	Major	Catastrophe	
Notifying of action at hand	47 (41.22)	2 (4.25)	22 (46.80)	16 (34.04)	7 (14.89)	2.16 ± 0.80
Existence of necessary controls	15 (13.15)	0	11 (73.33)	4 (26.66)	0	1.92 ± 0.50
Visibility and clarity of controls for users	27 (23.68)	0	10 (37.03)	16 (59.25)	1 (3.70)	2.21 ± 0.53
User perception of the work process in the system	25 (21.92)	1 (4)	10 (40)	11 (44)	3 (12)	2.17 ± 0.74
Total		3 (2.63)	53 (46.49)	47 (41.22)	11 (9.64)	

Table 3 Categorizing usability problems by tasks

Sub-task	n (%)	Severity n (%)				Average severity
		Cosmetic	Minor	Major	Catastrophe	
Authentication	1 (9.64)	0	0	1 (100)	0	2.56 ± 0.09
Recording patient identity information	38 (33.33)	3 (7.89)	23 (60.52)	9 (23.68)	3 (7.89)	1.87 ± 0.71
Recording patient tumor specification	29 (25.43)	0	11 (37.93)	16 (55.17)	2 (6.89)	2.28 ± 0.67
Removing patient duplications	41 (35.96)	0	18 (43.90)	17 (41.46)	6 (14.63)	2.27 ± 0.66
Reporting patients information	2 (1.75)	0	0	2 (100)	0	2.69 ± 0.26
System settings and exit	3 (2.63)	0	2 (66.66)	1 (33.33)	0	0.61 ± 2.21
Total		3 (2.63)	54 (47.36)	46 (40.35)	11 (9.64)	

Table 4 Demographic information of the participants

Demographic information	Number (%)
<i>Age</i>	
< 30	9 (33.33)
30–39	6 (22.22)
40–49	9 (33.33)
>= 50	3 (11.11)
<i>Gender</i>	
Male	8 (29.62)
Female	19 (70.37)
<i>Education degree</i>	
Bachelor's	14 (51.85)
Master's and higher	13 (48.14)
<i>Computer-related education degree</i>	
Yes	15 (55.55)
No	12 (44.44)
<i>Computer skills</i>	
Elementary or intermediate	10 (37.03)
Advanced	17 (62.96)
<i>Computer use</i>	
Daily	10 (37)
Two to several times a week	9 (33.3)
Two to several times a year	4 (14.9)
Having no access to computer, except at working place	4 (14.8)
<i>Internet use</i>	
Two to several hours a day	23 (85.2)
Less than one hour a day	4 (14.8)
<i>Work experience in cancer registry</i>	
Less than a year	3 (11.1)
Between 1 and 2 years	9 (33.3)
Between 2 and 10 years	12 (44.4)
More than 10 years	3 (11.1)
<i>Experience with the cancer registry system</i>	
Less than 6 months	4 (14.8)
Between 6 months to 1 year	2 (7.4)
Between 1 and 2 years	7 (25.9)
More than 2 years	14 (51.9)
<i>History of using other health information systems</i>	
Yes	24 (88.9)
No	3 (11.1)

registration unit of a Medical Sciences University. More than half of the participants (n = 14) stated that they have one to two years of experience using the cancer registry system. Most of the participants (n = 24) also had a history of using other health information systems.

The users' agreement with the usability problems identified by CW was CI 95% = 0.9 (0.96, 0.98). The most severe problems were related to the "Reporting patients information" (2.48) in the system. Problems concerning

"Authentication" of a patient (1.85) and "Recording patient identity information" including manual entry of patient's date of birth, displaying inappropriate format for patient's date of birth after manual entry, the impossibility of manual entry of characteristics such as nationality, province of birth, and improper system alert for changing the language (1.88) had the lowest severity (Fig. 8).

Discussion

Principal findings

In this study, evaluators identified 114 usability problems related to the design of a cancer registry system. From the evaluators' point of view, in categorizing problems based on CW questions, most of the problems were related to the fact that a user can not set a goal at each stage of working with the system. Mismatch of the system controls with the action at hand, and unclear feedback about what the system is doing, each accounted for about a quarter of the problems. In the categorization of problems based on user tasks, about three-quarters of the problems concerned the tasks of "removing patient duplications" and "recording patient identity information". Twenty-seven users participated in the second phase of the study in order to check their agreement with the usability problems. These users strongly agreed with the problems identified by CW.

Evaluation of the cancer registry system using CW

In this study, eight evaluators identified a large number of system usability problems in a CW evaluation. This result is in line with the results of other studies indicating that CW can identify a large number of usability problems in an information system with 3–8 evaluators [6, 20, 41].

In categorizing problems based on CW questions, most of the usability problems emerged since a user could not set a goal at each stage of working with the system. This result is in line with the results of a study [20] that used CW to evaluate a maternity information system and showed that less than a quarter of the problems are related to the lack of knowledge concerning the action at hand at every stage of working with the system. One of the reasons for such problems is the lack of instructions on using the system or the lack of clues to guide the user through the system, which makes it difficult to find a way to work with the system, imposes a cognitive burden on their memory, and leads to fatigue and dissatisfaction of users with information systems [42].

In this study, a quarter of the problems arose due to the inconsistency of system controls such as icons, buttons, and menus with their intended actions. This result confirms the results of a previous study evaluating a remote drug delivery system by heuristic and

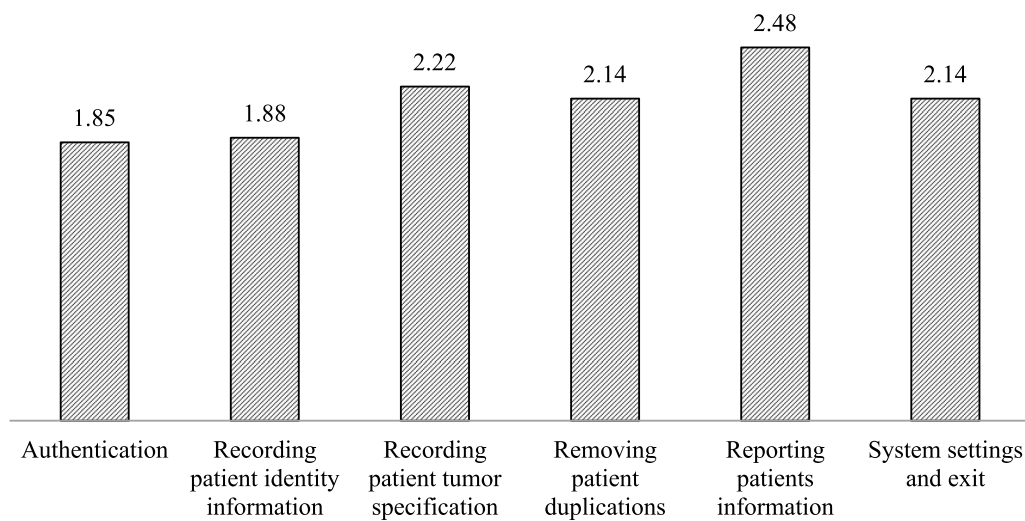


Fig. 8 The severity of usability problems from the users' point of view

CW evaluations, which showed that it is difficult for a user to find the relationship between system controls and their corresponding actions. For example, unclear categorization of the user interface elements confuses users on how to navigate the available options. According to other studies [43–46], problems such as small size and closeness of icons and buttons, inappropriate labels, small and inconsistent fonts, irrelevant colors, vague and obscure menus, and Improper placement of controls are among design problems that prevent users from understanding the relationship between system controls and corresponding actions. Ideally, the design of the controls in a system's interface should match the users' mental model to facilitate users' interpretation of the controls' functions. Therefore, system designers can minimize users' memory load by using meaningful and understandable symbols on the interface [47, 48].

According to studies [20, 22, 49], CW can identify problems related to the users' understanding of system performance. In this regard, almost a quarter of the problems identified in the present study showed that the users do not have a proper understanding of what the system does at different stages. A previous study [49] revealed that poor system response, vague feedback or lack of feedback, and even delays in receiving feedback result in a difficult understanding of system performance. Therefore, providing appropriate audio and video feedback after each user's action prevents confusion and user errors, and ultimately improves performance.

Categorization of problems according to tasks showed that most problems were associated with the tasks having a higher number of actions compared to other tasks. The

highest number of problems in the cancer registry system evaluated in this study occurred during "recording patient identity" such as name, surname, age, and address of residence; as well as during the process of "removing patient duplications". This result confirms the results of two former studies that evaluated a drug information system [9], and a health research system [50]. These studies have shown that tasks that require more action raise more problems for users. Hence, these tasks increase the cognitive load of the user's memory and require more time to be completed. Thus, it is recommended to reduce the number of actions required to perform tasks in the development and redesign of the system.

Agreement of the cancer registry users with the problems identified by CW

In expert-based methods, experts primarily identify general problems related to the user interface. But a user-testing method only identifies problems that users encounter when performing their tasks [51]. Also, expert-based methods identify a higher number of problems compared to user-testing methods [51, 52]. The agreement of the cancer registry users with the problems identified by CW was used to investigate to what extent users confirm the issues identified by CW. According to the results, users have a high agreement with the problems identified by CW. In previous studies, the users' agreement with the problems identified by CW has not been measured. The only paper [17] that addressed the agreement of users with the results of such studies has shown that users strongly agree with the problems identified by Heuristic Evaluation. Other relevant studies measured the agreement between the results of the CW

method and the results of the other user-centered evaluation methods. These studies that evaluated self-care mobile applications for patients with diabetes [24] and Acquired Immune Deficiency Syndrome (AIDS) [23] have shown that CW and think-aloud identify similar problems and users and evaluators strongly agree on the required user interface modifications such as changes in color, font, and text size. Besides, our study confirms the results of studies that measured the agreement between the results of other expert-centered evaluation methods and the results of user-centered evaluation methods [25, 26]. These studies that evaluated a virtual reality program on eating behavior [26] and the self-care system of AIDS patients [25] have reported a high agreement between the results of the Heuristic Evaluation and user-centered evaluation methods.

Study implications

The results of this study and previous studies have shown that the lack of a help function for using the system or the lack of clues to guide the user confuses users about what to do at each stage of working with the system. Therefore, to avoid user confusion, system designers and developers embed the help functionality in a suitable place in the system and use clues to facilitate navigation throughout the system. When the user could not associate the intended actions to their controls in the system, this usability problem imposes a heavy cognitive load on the user's memory. On the other hand, poor system response and lack of appropriate and timely feedback prohibit understanding the performance of the system. Hence, system developers should design meaningful and understandable controls based on users' mental models and design and provide appropriate audio and video feedback for each user action, to help the user understand what the system is doing.

According to the results of this study, most problems originate from tasks that require a higher number of actions. Tasks with several actions in an information system entail long-term interaction with the system or moving between different pages, which impose a heavier cognitive burden on the user's working memory [53]. Therefore, it is recommended to develop systems that allow users to perform their tasks by taking fewer actions.

One of the challenges of task-based evaluation methods such as CW is the inappropriate selection of tasks in the preparation phase of the evaluation. Inappropriate selection of tasks may prevent the identification of important and frequently encountered usability problems [28]. Usability problems, especially the most frequent ones, increase users' cognitive load and lead to fatigue and burnout [54]. Therefore, identifying the recurring problems helps improve the design of information systems

and increases system performance by reducing the cognitive burden and reducing the learning time for users. In this study, we evaluated a cancer registry system by selecting the most routine and frequent tasks performed by users. Therefore, the problems identified in this study are the most important problems of the system from the users' point of view in the real environment.

Due to the large number of problems identified by CW in this study, we measured the degree of users' agreement solely with more severe problems. Our results indicate a high degree of user agreement with the identified problems. These results are in line with the results of a study [17] that evaluated an emergency information system and showed that the level of user agreement with the problems identified by Heuristic Evaluation is high. Therefore, whenever access to real users is difficult, expert-based evaluation methods such as CW can be used effectively to evaluate information systems and find usability problems.

Study limitations and future studies

This study had four limitations. First, The main objective of this study was to assess the agreement rate of users with usability problems identified by the CW method. Therefore, some of the usability problems that users may encounter in a real environment could be overlooked. Future studies can identify all the system usability problems faced by users through user testing methods. Second, participating users were provided only with the usability problems identified by CW. Subsequent studies can use open-ended questions to determine whether users add other problems not identified by CW. Third, to increase the accuracy of answers and response rate, from the large number of problems identified in the CW evaluation, only the problems with higher severities were provided to users. Studies with a lower number of problems, can measure user agreement with all problems. Fourth, we invited all the users of the cancer registry system in Iran to participate in this study. Despite sending several reminders to users to participate in the study, the response rate was not high. Examining the agreement of a higher number of users with usability problems can increase the accuracy of the results. However, this is the first study examining the agreement of users with the results of a CW of a healthcare information system and the results can be used for evaluating similar systems.

Conclusion

In categorizing problems based on CW questions, most of the system problems are related to the issue that the user cannot set a goal at each stage of working with the system. Therefore, system designers and developers should provide a help function in a convenient location

of the interface and use informational clues to facilitate the navigation of users throughout the system. In addition, half of the problems indicated that the user cannot make an accurate connection between his goal and the system controls, or he does not have a proper understanding of what the system is doing at different stages. Therefore, to prevent user confusion, controls should be used that match user's mental model, and the system should provide appropriate feedback after each user action to help the user understand what the system is doing. Since most problems occur when completing tasks with a higher number of actions, minimizing the number of actions in each task improves the design and usability. In this study, the scenarios were developed based on the most routine and frequent tasks of users. Therefore, the results can be used for designing systems having similar tasks. Due to the high agreement of users with the problems identified by CW, the results of this study indicate that in the absence of users, the CW can detect a large number of problems, especially important problems that may be encountered by users in the real environment.

Abbreviations

CW	Cognitive Walkthrough
SPSS	Statistical Package for Social Science
SMS	Short Message Service
AIDS	Acquired Immune Deficiency Syndrome
KUMS	Kerman University of Medical Sciences

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12911-023-02120-8>.

Additional file 1. Evaluation checklist.

Additional file 2. Problems form.

Additional file 3. Problems severity form.

Additional file 4. Online questionnaire.

Additional file 5. Identified problems.

Acknowledgements

The authors would like to thank the participants in the study

Author contributions

FB and FA and RK contributed to the conception and design of the study, acquisition, analysis and interpretation of the data, and drafting the paper. MS contributed to the collect data and increase of the sample size in revised version. All authors read and approved the final version of the article submitted.

Funding

No funding was obtained for this study.

Availability of data and materials

The data generated and analyzed during this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

In this study, we did not collect identifiable personal data. The questionnaires were completely filled out without names, surnames, and other identifiable data. Therefore, to participate in the study, oral informed consent was obtained from the participants, and the participants were assured that their information would be kept confidential. This study was approved by the research ethics committee of KUMS (ethics code: IR.KMU.REC.1399.41). In this study, all methods were performed in accordance with the relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 6 May 2022 Accepted: 23 January 2023

Published online: 30 January 2023

References

- World Health Organization (WHO). Cancer. 2019. <https://www.who.int/cancer/en/>. Accessed 22 Dec 2022.
- Valsecchi MG, Steliarova-Foucher E. Cancer registration in developing countries: luxury or necessity? *Lancet Oncol*. 2008;9(2):159–67.
- Sobue T. Cancer registration system: an introduction. *Int J Clin Oncol*. 2008;13(2):89.
- Choi J, Bakken S. Web-based education for low-literate parents in Neonatal Intensive Care Unit: development of a website and heuristic evaluation and usability testing. *Int J Med Informatics*. 2010;79(8):565–75.
- Sutcliffe A. Multimedia and virtual reality: designing multisensory user interfaces. 1st ed. New York: Psychology Press; 2003.
- Khajouei R, Zahirri Esfahani M, Jahani Y. Comparison of heuristic and cognitive walkthrough usability evaluation methods for evaluating health information systems. *J Am Med Inform Assoc*. 2017;24(e1):e55–60.
- Simon SR, Kaushal R, Cleary PD, Jenter CA, Volk LA, Poon EG, et al. Correlates of electronic health record adoption in office practices: a statewide survey. *J Am Med Inform Assoc*. 2007;14(1):110–7.
- Nielsen J. Finding usability problems through heuristic evaluation. In: Proceedings of the SIGCHI conference on human factors in computing systems; Monterey, California, USA: ASSOCIATION for Computing Machinery; 1992. DOI:<https://doi.org/10.1145/142750.142834>.
- Yasemi Z, Rahimi B, Khajouei R, Yusefzadeh H. Evaluating the usability of a nationwide pharmacy information system in Iran: application of Nielsen's heuristics. *J Clin Res Paramed Sci*. 2018. <https://doi.org/10.5812/jcrps.80331>.
- ISO. ISO Online browsing platform. In: Switzerland. 2018. <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>. Accessed 12 Dec 2022.
- Sibanda M, Summers R, Meyer JC. A comparison of five international clinical trial registers with the South African register for access to information and usability. *Pan Afr Med J*. 2018;29:224.
- Nielsen J. Usability engineering. Amsterdam: Elsevier Science; 1994.
- Nielsen J, editor. Usability engineering at a discount. In: Proceedings of the third international conference on human-computer interaction on Designing and using human-computer interfaces and knowledge based systems (2nd ed); 1989.
- Georgsson M, Staggers N, Årsand E, Kushniruk A. Employing a user-centered cognitive walkthrough to evaluate a mHealth diabetes self-management application: a case study and beginning method validation. *J Biomed Inform*. 2019;91:103–10.
- Lewis C, Wharton C. Cognitive walkthroughs. In: Handbook of human-computer interaction (1st ed). Amsterdam: Elsevier; 1997.
- Bligård L-O, Osvalder A-L. Enhanced cognitive walkthrough: development of the cognitive walkthrough method to better predict, identify, and present usability problems. *Adv Hum-Comput Interact*. 2013. <https://doi.org/10.1155/2013/931698>.

17. Khajouei R, Ameri A, Jahani Y. Evaluating the agreement of users with usability problems identified by heuristic evaluation. *Int J Med Inform.* 2018;117:13–8.
18. Cohen T, Kaufman D, White T, Segal G, Staub AB, Patel V, et al. Cognitive evaluation of an innovative psychiatric clinical knowledge enhancement system. *Stud Health Technol Inform.* 2004;107(2):1295–9.
19. Usselman E, Borycki EM, Kushniruk AW. The evaluation of electronic perioperative nursing documentation using a cognitive walkthrough approach. *Stud Health Technol Inform.* 2015;208:331–6.
20. Ghalibaf AK, Jangi M, Habibi MRM, Zangouei S, Khajouei R. Usability evaluation of obstetrics and gynecology information system using cognitive walkthrough method. *Electron Physician.* 2018;10(4):6682–8.
21. Sears A. Heuristic walkthroughs: finding the problems without the noise. *Int J Hum-Comput Interact.* 1997;9(3):213–34.
22. Farzandipour M, Nabovati E, Sadeqi JM. Comparison of usability evaluation methods for a health information system: heuristic evaluation versus cognitive walkthrough method. *BMC Med Inform Decis Mak.* 2022;22(1):1–11.
23. Beauchemin M, Gradilla M, Baik D, Cho H, Schnall R. A multi-step usability evaluation of a self-management app to support medication adherence in persons living with HIV. *Int J Med Inform.* 2019;122:37–44.
24. Georgsson M. NASA RTLX as a novel assessment for determining cognitive load and user acceptance of expert and user-based evaluation methods exemplified through a mHealth diabetes self-management application evaluation. *Stud Health Technol Inform.* 2019;261:185–90.
25. Lai T-Y. Iterative refinement of a tailored system for self-care management of depressive symptoms in people living with HIV/AIDS through heuristic evaluation and end user testing. *Int J Med Inform.* 2007;76:17–24.
26. Langlet BS, Odegi D, Zandian M, Nilstam J, Södersten P, Bergh C. Virtual reality app for treating eating behavior in eating disorders: development and usability study. *JMIR Serious Games.* 2021;9(2):e24998.
27. Polson PG, Lewis CH. Theory-based design for easily learned interfaces. *Hum-Comput Interact.* 1990;5(2):191–220.
28. Polson PG, Lewis C, Rieman J, Wharton C. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *Int J Man Mach Stud.* 1992;36(5):741–73.
29. Schmegg M. Sample size in usability studies. *Commun ACM.* 2012;55(4):64–70.
30. Philips M. Heuristic analysis for UX: How to run a usability evaluation. 2010–22 <https://www.toptal.com/designers/usability-testing/usability-analysis-how-to-run-a-heuristic-evaluation>. Accessed 12 Dec 2022.
31. Borsci S, Macredie RD, Barnett J, Martin J, Kuljis J, Young T. Reviewing and extending the five-user assumption: a grounded procedure for interaction evaluation. *ACM Trans Comput-Hum Interact (TOCHI).* 2013;20(5):1–23.
32. Shliachkov V. Sample size for usability study. Part 1. About Nielsen and probability. 2019. <https://uxplanet.org/sample-size-for-usability-study-part-1-about-nielsen-and-probability-effecdfbfa95>. Accessed 12 Dec 2022.
33. Patterson ES, Zhang J, Abbott P, Gibbons MC, Lowry SZ, Quinn MT, et al. Enhancing electronic health record usability in pediatric patient care: a scenario-based approach. *Jt Comm J Qual Patient Safety.* 2013;39(3):129–35.
34. Nielsen J. Usability inspection methods. In: Conference companion on human factors in computing systems. ConferenceCompan. 1994; 413–4.
35. Travis D. How to prioritise usability problems. 2009. <http://www.userfocus.co.uk/articles/prioritise.html>. Accessed 12 Dec 2022.
36. Welbie M, Wittink H, Westerman MJ, Topper I, Snoei J, Devillé WL. A mobile patient-reported outcome measure app with talking touch-screen: usability assessment. *JMIR formative research.* 2019;3(3):e11617.
37. Khajouei R, Peute LW, Hasman A, Jaspers MW. Classification and prioritization of usability problems using an augmented classification scheme. *J Biomed Inform.* 2011;44(6):948–57.
38. Karahoca A, Bayraktar E, Tatoglu E, Karahoca D. Information system design for a hospital emergency department: a usability analysis of software prototypes. *J Biomed Inform.* 2010;43(2):224–32.
39. Brown B, Balatsoukas P, Williams R, Sperrin M, Buchan I. Interface design recommendations for computerised clinical audit and feedback: hybrid usability evidence from a research-led system. *Int J Med Informatics.* 2016;94:191–206.
40. Nielsen J. Severity ratings for usability problems. 1994. <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>. Accessed 12 Dec 2022.
41. Jadhav D, Bhutkar G, Mehta V. Usability evaluation of messenger applications for Android phones using cognitive walkthrough. In: Proceedings of the 11th Asia pacific conference on computer human interaction. 2013; 9–11.
42. Horsky J, Kaufman DR, Oppenheim MI, Patel VL. A framework for analyzing the cognitive complexity of computer-assisted clinical ordering. *J Biomed Inform.* 2003;36(1–2):4–22.
43. Guo J, Iribarren S, Kapsandoy S, Perri S, Staggers N. Usability evaluation of an electronic medication administration record (eMAR) application. *Appl Clin Inform.* 2011;2(2):202–24.
44. Gatsou C, Politis A, Zevgolis D. the importance of mobile interface icons on user interaction. *Int J Comput Sci Appl.* 2012;9:92–107.
45. Zaini NA, Fadzilah S, Siti T. Evaluation of API interface design by applying cognitive walkthrough. *Int J Adv Comput Sci Appl.* 2019;10:306–15.
46. Abdulai A-F, Howard AF, Noga H, Yong PJ, Currie LM. Application of Anti-Stigma design heuristics for usability inspection. *Stud Health Technol Informatics.* 2021;284:239–43.
47. Farzandipour M, Nabovati E, Tadayon H, Sadeqi JM. Usability evaluation of a nursing information system by applying cognitive walkthrough method. *Int J Med Informatics.* 2021. <https://doi.org/10.1016/j.ijmedinf.2021.104459>.
48. Holden RJ, Campbell NL, Abebe E, Clark DO, Ferguson D, Bodke K, et al. Usability and feasibility of consumer-facing technology to reduce unsafe medication use by older adults. *Res Social Adm Pharm.* 2020;16(11):54–61.
49. Ligans FM, Romagnoli KM, Browell S, Hochheiser HS, Handler SM. Assessing the usability of a telemedicine-based medication delivery unit for older adults through inspection methods. USA: American Medical Informatics Association; 2011. p. 795–804.
50. Di Wang DRK, Mendonca EA, Seol Y-H, Johnson SB, Cimino JJ, editors. The cognitive demands of an innovative query user interface. USA: American Medical Informatics Association; 2002. p. 850–4.
51. Yen P-Y, Bakken S, editors. A comparison of usability evaluation methods: heuristic evaluation versus end-user think-aloud protocol—an example from a web-based communication tool for nurse scheduling. USA: American Medical Informatics Association; 2009. p. 714–8.
52. Maguire M, Isherwood P, editors. A comparison of user testing and heuristic evaluation methods for identifying website usability problems. In: International conference of design, user experience, and usability; 2018; 409–29.
53. Kaufman DR, Starren J, Patel VL, Morin PC, Hilliman C, Pevzner J, et al. A cognitive framework for understanding barriers to the productive use of a diabetes home telemedicine system. *AMIA Annu Symp Proc AMIA Symp.* 2003;2003:356–60.
54. Gude WT, Simon AC, Peute LW, Holleman F, Hoekstra JB, Peek N, et al. Formative usability evaluation of a web-based insulin self-titration system: preliminary results. *Stud Health Technol Inform.* 2012;180:1209–11.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

