Exploring usability problems of mHealth applications developed for cervical cancer: An empirical study

Md Abdur Razzak[1], Muhammad Nazrul Islam[2] (iD), Tanjila Broti[2], Elma Shumsun Kamal[2] and Samia Zahan[2]

Abstract

Objectives: Nowadays, mobile health applications are developed to raise awareness and facilitate screening and treatment of cervical cancer, while a very few studies have been conducted focusing on the measurement and assurance of usability and exploring the acceptable user experience of such applications. Usability issues become a crucial concern for such cervical-cancer-related applications because users with diverse backgrounds in terms of education, information technology literacy, and geographic reasons are required to access those applications. The objective of this research is to evaluate the usability of mobile health applications developed for cervical cancer patients.

Methods: Two evaluation studies were conducted following the expert evaluation and a questionnaire-based user study. A total of four cervical-cancer-related applications that are focusing on the *Awareness* and *Diagnosis* theme were selected and each of the applications was evaluated by four usability experts. Then, a user study ($n=80$) based on the Goal Question Metric was conducted to reveal the usability problems of four selected applications. Finally, findings of both evaluations were aggregated and analyzed.

Results: Both approaches showed that all applications suffer from several usability problems while "Cervical Cancer Guide" performs better and "Cervical Cancer Tracker" showed the least in performance from the usability perspective. Again, the Goal Question Metric performs noticeably better in assessing the learnability of the applications, while the analytical heuristic evaluation performs better in identifying the issues that cause user annoyance.

Conclusion: The methodology adopted and the usability problems revealed through this study can be well utilized by the information technology professionals or user interface designers for designing, evaluating, and developing the cervical-cancer-related applications with enhanced usability and user experience.

Keywords

Cervical cancer, usability evaluation, heuristic evaluation, mobile application, health informatics, Goal Question Metric

Introduction

Cervical cancer is the fourth most common cancer among women globally,[1] with an estimated 604,000 new cases and 342,000 deaths in 2020.[2,3] It is the most common gynecologic cancers including cervical, uterine, ovarian, and vaginal and vulvar. Nearly in all countries, a huge number of populations (around 20%) remain out of screening programs which later cause cervical cancer.[4] Cervical cancer has an enormous socioeconomic impact on patients in terms of social discrimination, loss of body image, loss of sexual functioning, loss of femininity, loss of income, financial distress, and work and employment challenges.[5] Again, though

breast cancer is the second leading cause of cancer death in women, but the decrease death rates in breast cancer is believed to be the result of finding breast cancer earlier

[1]Faculty of Science and Technology, Bangladesh University of Professionals (BUP), Dhaka, Bangladesh
[2]Department of Computer Science and Engineering, Military Institute of Science and Technology (MIST), Dhaka, Bangladesh

Corresponding author:
Muhammad Nazrul Islam, Department of Computer Science and Engineering, Military Institute of Science and Technology (MIST), Mirpur Cantonment, Dhaka 1216, Bangladesh.
Email: nazrul@cse.mist.ac.bd

through screening and increased awareness, and by providing better treatments. Similarly, a noticeable number of studies were conducted focusing on the development and usability evaluation of mobile applications for breast cancer.[6–8] On the contrary, cervical cancer involves more sensitive organs of the women's body and has not yet reached a satisfactory level of public awareness. The recent progressive infection rate, the social hood of ignorance, and the crying necessity of raising awareness are the key motivations behind choosing cervical cancer domain as the concentrating subject for this research.

The World Health Assembly adopted the global strategy for cervical cancer elimination, considering it as a public health problem. This health issue can be effectively treated to cure if prognosis is done at its emerging stage by regular cervical cancer screenings, which include a Pap test and an human papillomavirus (HPV) test.[9] However, a number of studies have also been conducted to explore and evaluate the related scopes.[10–12] Mobile applications as well as online platforms can help to detect and prevent cervical cancer in its premature phase. For example, in Muljo et al.,[13] a mobile is developed following the analysis, design, development, implementation, evaluation approach for early detection of cervical cancer in Indonesia. Similarly, to increase the practices of cervical cancer screening among the vulnerable community, Lee et al.[14] developed an application to encourage Korean ethnic females to get a Pap test (in the United States) and found that the persuasiveness of mobile apps enriching learning about cervical cancer and influencing screening. Likewise, Stocks et al.[15] developed an mHealth application to assist HPV-based screening in western Kenya, while Stocks et al.[15] showed that the mHealth app was useful for community health volunteers to conduct HPV-dependent screening of cervical cancer and facilitated for answering frequently asked questions and providing counseling. Again, the fast-growing mobile technologies create awareness for HPV infection, vaccination, and screening programs among women, especially younger ones in an efficient manner. For example, Ruiz-Lopez et al.[4] developed a learning app named FightHPV for creating awareness regarding the prevention and risk factors related to HPV among the mobile technology user. The study found that the FightHPV was successful in increasing user knowledge and providing a positive influence against HPV. In another study, Quercia et al.[9] explored the uses of a mHealth application to collect data related to cervical cancer screening campaigns for monitoring the women participation in such campaigns. As such, cervical-cancer-related mobile apps can provide a solution for the detection,[10] classification,[12] and screening[16] of cervix cancer and creation of awareness among people to diminish health disparity through learning apps.[17,18] Moreover, the use of mobile health applications may provide betterment of the safety and gratification of the patient retaining the up-to-date health information.[19,20]

Nowadays, the uses of machine learning (ML) techniques notably grow in the detection of different risk factors for different diseases among patients.[21–23] As such, instead of using the traditional prognosis of recurrent cervical cancer, several studies have been conducted focusing prediction, screening, and diagnosis using the ML. For example, Tseng et al.[24] explored three ML techniques named support vector machine, C5.0, and extreme learning machine to detect the risk factors associated with recurrent cervical cancer. They found that the C5.0 classifier is most suitable for detecting the recurrence risk factors with an average categorizing rate of 96.00%.

Usability is a key quality attribute for any mobile applications as it provides the ease of use for specific users.[25–27] The ease of use of the mobile application is deeply rooted in the effectiveness, satisfaction, and efficiency of the application.[28,29] Although several studies have been conducted as well as several number of applications have been developed on cervical cancer domain till now, their use is not still making much visible impact on users' life as there are some lacking factors related to usability, user experience (UX) and user interface (UI) design, etc., which are subject to analysis. As such the prior studies disclose the necessity of evaluating such mobile applications. In this regard, Hussain and Ferneley[30] suggested that heuristic evaluation (HE) would be more concise if the checklist can be mapped to some metric for assessing the usability, while Caldiera and Rombachin[31] suggested to conduct Goal Question Metric (GQM) study along with a historically used approach (e.g., HE) for improving the assessment quality. However, among the stated studies, only a few studies[17] are dedicated to evaluating multiple cervical cancer mobile apps, while most of the applications were developed for Thai native users and a less focus has been given to evaluate the system through heuristic evaluation.

Thus, the objective of this research is to assess the usability of mobile health applications developed for cervical cancer. To achieve this research objective, four cervical-cancer-related applications, thematically focused on "awareness" and "diagnosis," were selected and evaluated through two approaches, that includes, GQM[30] and HE.[32] Finally, the findings of both approaches were analyzed and compared.

Methods

From a methodological perspective, this study followed an explorative research approach where two evaluation approaches (heuristic evaluation and GQM0) were adopted to reveal the usability problems of cervical-cancer-related applications (see Figure 1). As such, both qualitative and quantitative data were collected. So, the exact nature of the article could be considered as mixed method.

Choosing the related applications

We searched for mobile applications related to cervical cancer at this stage. First, a string-based search[17] was carried out
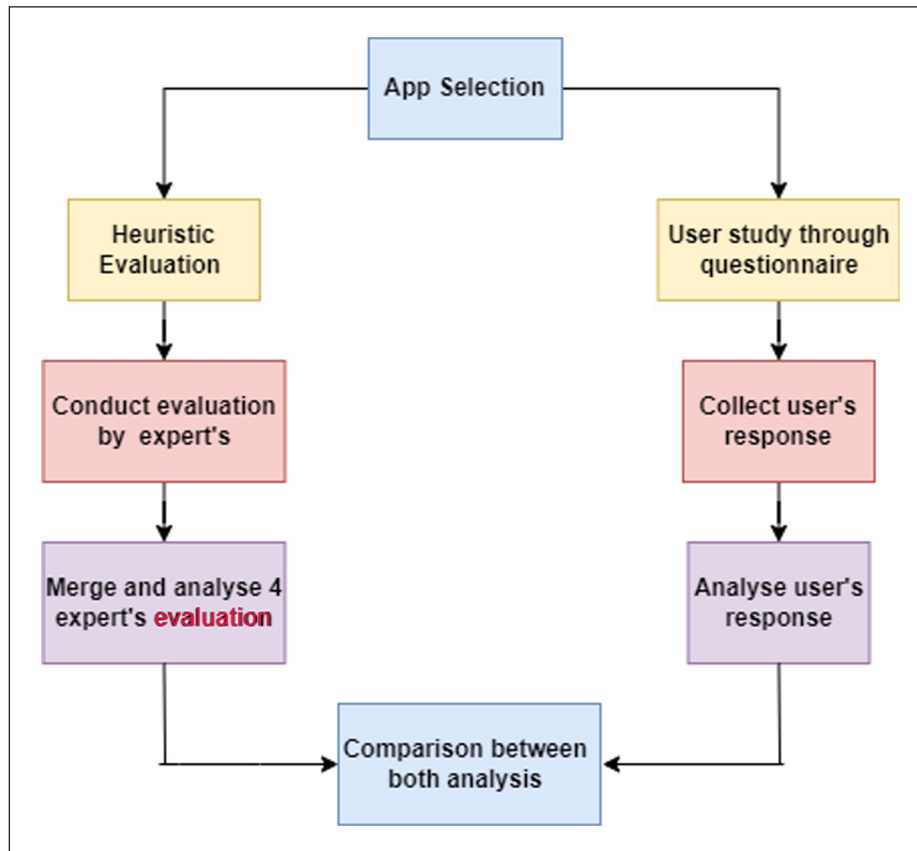
Figure 1. Flow diagram of the process.

in Google play store during October to November 2022 with the strings like "cervical cancer," "cervical cancer awareness," and "cervical cancer screening." Second, applications found as a search result were examined carefully to make sure that it is for cervical cancer. Third, several apps were selected which are highly rated (3+) and thematically focused on awareness and diagnosis, while carefully handling all duplicate discoveries throughout the above steps. Finally, the *Cervical Cancer Guide* and *Cervical Cancer Forum* applications that are focusing on the "Awareness" theme and the *Cervical Cancer Tracker* and *Figo Staging* applications from the "Diagnosis" theme were selected for evaluation. Hereafter, the selected apps namely *Cervical Cancer Guide, Cervical Cancer Forum, Cervical Cancer Tracker*, and *Figo Staging* are referred to as App1, App2, App3, and App4, respectively. The *Cervical Cancer Guide (App1)* is an information-based application containing guidelines regarding cervical cancer, while the *Cervical Cancer Forum (App2)* provides users a discord-like platform where they can get connected with doctors and other patients and can communicate in one-to-one manners or in a session. *The Cervical Cancer Tracker (App3)* provides static information to track cervical cancer and provides interactive information. The last one, *Figo Staging (App4)* contains information

about the stages of cervical cancer, the process and ways to determine cancer stages.

Evaluations

The selected apps were evaluated following the heuristic evaluation[33–35] and the user-based evaluation.[36,37] For heuristic evaluation, Nielsen's 10 Heuristics[38] were used (see Table 1) and Nielson's severity ranking scale (0–4) was adopted for rating the severity of the revealed issues. The evaluation was performed by the four experts/authors (three females and one male) who graduated in computer science and conducted several courses related to human–computer interaction, usability evaluation, and interface design. All of them have 2–6 years of professional experience and evaluated a minimum of 4–6 mobile and web applications. At first, each evaluator evaluated each of the selected applications separately. Finally, their findings were aggregated to enumerate the usability problems.

The study also adopted the GQM approach[36,37] for usability evaluation. Here, a set of questionnaires was formulated following the GQM architecture (see Figure 2) and by adopting the parameters used for usability evaluation (e.g., user feedback, UX, design principles). The GQM approach

Table 1. Nielsen's 10 heuristics.

Heuristics	Questions
H1	Visibility of system status
H2	Match between system and the real world
H3	User control and freedom
H4	Consistency and standards
H5	Error prevention
H6	Recognition rather than recall
H7	Flexibility and efficiency of use
H8	Aesthetic and minimalist design
H9	Help users recognize, diagnose, and recover from errors
H10	Help and documentation
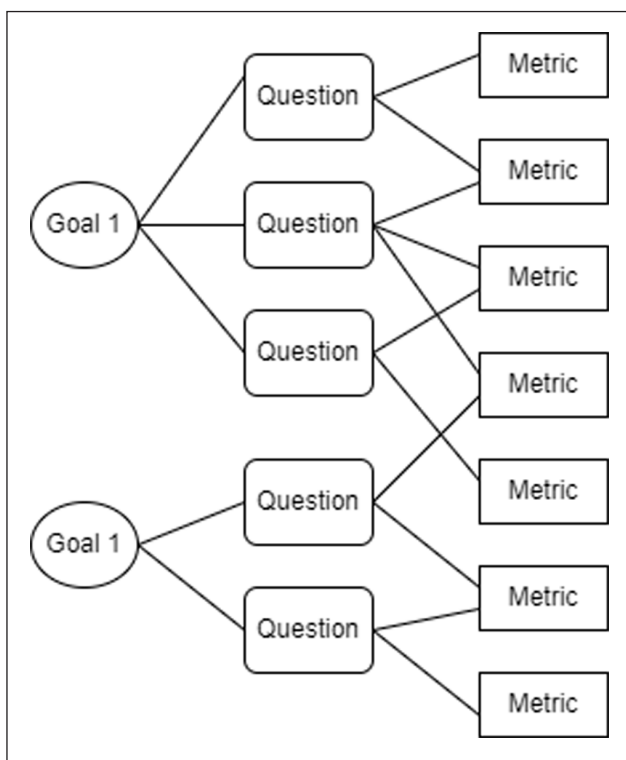
Figure 2. The architecture of GQM approach.[30]

defines goals first, goals are refined into questions and, subsequently, metrics are defined that may provide the information to answer these questions. By answering the questions, the measured data can be analyzed to identify if the goals are attained. Thus, using GQM, metrics are defined from a top-down perspective and analyzed bottom-up.[6] The hierarchical structure of GQM adopted from Hussain A and Ferneley[30] is shown in Figure 2.

In GQM approach (see Figure 2), first six goals were defined including (G1) to obtain user feedback on apps features; (G2) to understand the UX on the app usages; (G3) to learn the comparative market values with respect to the

existing other apps; (G4) to measure how the design principles are followed to design the selected apps; (G5) to know the ability of preventing error due to users' incorrect intervention; and (G6) to acquire the overall feedback of the selected applications. Second, a set of questions were prepared for each goal. Finally, one metric is proposed against each question except question number 17 where three metrics have been proposed. The selected goals, questions, and metrics of GQM approach are presented in Table 2.

The GQM-based questionnaires in Google form were distributed to the female students of the authors' institute and through social media. Data collection was carried out over 2 weeks and a total of 80 responses (see Table 3) were received and analyzed to reveal and assess the usability problems. All the respondents were female and had a good familiarity with the usage of internet, computer, and mobile applications. Their age varied from 20 to 57 with an average age 39. Here, all the respondents were aware of cervical cancer, while 15 respondents had good knowledge on this disease, since either them or any of their family members were affected by cervical cancer. Only 10 participants used cervical-cancer-related applications before, while none of them had used any of the selected applications. Again, 45% of the respondents were students and the rest 65% included teachers, lawyers, housewives, doctors, etc. No relationship was established prior with the respondents to study commencement and participants were informed the purpose of this research, the subject heading, and introductory brief written on the top of the question form.

To collect and analysis the study data, data saturation for the qualitative data was also considered. For the heuristic evaluation, four experts independently evaluated each selected mobile applications and then aggregated their findings for each applications and for all applications. Evaluating the application by more experts may not raveled any new usability problems and would not be cost and time efficient. Again, only few qualitative data were collected through GQM and responses from 80 users were considered as data saturation. For example, responded say about what other features they expect from a specific cervical-cancer-related mobile app. In response, we received few features repetitively from the respondents.

Results

Findings through heuristics evaluation

Selected apps were evaluated through HE.[7] A few examples of HE are presented here. Please see Table 1 for the mapping of heuristics. In App3 (Cervical Cancer Tracker), inconsistency and mismatch observed between the system and real world (see Figure 3). While pressing the "Home" option from the dashboard, it unexpectedly redirects to a new page with a heading, named "Information"! Again, clicking on the icon ("I") on the "Hospital page," it surprisingly redirects to

Table 2. The proposed goals, questions, and metrics of GQM approach.

Goals	Questions	Metrics
G1: User feedback on App's features	Q1: How much does the app help you to solve your problem/achieve your goal?	Goal achievement
	Q2: How attractive are the features of the app to you?	Attractiveness
	Q3: Did the features of the app work as expected?	Usefulness
	Q4: How would you feel if you can no longer use the app?	Emotional attachment
	Q5: What features would you like to add to the app?	Extension
G2: UX	Q6: Which app has the most satisfying design (UI) as per you?	Design (UI)
	Q7: Which app is the least likable to you?	Likeability
	Q8: What confused/annoyed you about the app?	Negativity
G3: Comparative market values	Q9: How did you get to know the app?	App reachability
	Q10: How does the app compare with competitors?	Competitiveness
	Q11: What are the alternatives that you are considering to the app?	Comparison
G4: Design principles	Q12: Is there a clearly identified link to the home page?	Visibility
	Q13: Is the system's icon or representation relevant to you?	Relevance
	Q14: Does the system look consistent (text, color) throughout the navigation?	Consistency
	Q15: How does the navigation look to you?	Documentation
	Q16: Does the system show easy search, limited number of buttons, and clear labeling/terminology?	Easy search, limited button/ navigation link, intuitive terminology
G5: Error management	Q17: Is the application alerting you for giving any wrong interaction/input?	Error prevention
G6: Overall feedback	Q18: How would you rate the overall quality of the app?	Quality
	Q19: How likely are you to recommend this app to a friend or colleague?	Recommendation
	Q20: Do you have any additional comments/suggestions?	Suggestion

Table 3. The participants profile.

Sex		Familiarity with the internet, computer and mobile apps	Age	Knowledge of cervical cancer		Familiarity with the studied apps		Profession	
Male	Female	100%	20–57	Good	Moderate	AppsUsed	Apps not-Used	Student	Others
0%	100%			15	65	10	70	45%	65%

the same "Information" page (which was shown on clicking "Home") that contains information about cervical cancer screening but no information about hospitals! Thus, this design violates H4 (Consistency and standards) and H2 (Match between system and the real world). Similarly, in the Figo Staging App (App4), there is a button labeled as "Click here for interactive staging" (see Figure 4). But on clicking this button instead of showing any interactive service it redirects to an ever-loading page. By not working consistently as promised, this violates H4. Moreover, the disturbing part is, there is no back button to get back to the main app from that ever-loading page. On clicking the default back button of the phone, it gets out of the application. As there was no user control and flexibility, it violated H3 (User control and freedom) and H7 (Flexibility and efficiency of use) with a severity level 4. Again, in the "Help" page of App4, the help section is empty which violates H10 (Help and documentation). Similarly, in App1 (Cervical cancer Guide), the "What is Cervical cancer?" page provides the abbreviated terms that may create difficulties to understand the clinical terms to end users, for example, HPV. As such, each app may have

multiple problems of heuristic violence. In Table 4, we have shown the summation of the number of problems of each heuristic violation and also their severity in an average count. The summary findings from four experts are presented in Table 4. The data showed that App4 has maximum number of problems and H8 violates most compared to all the selected apps and heuristics. The average severity score is shown highest in App4 while comparatively less score shown in App1.

Finding's through GQM approach

Finding of GQM approach (in percentage) is presented in Tables 5–7. For example, in response to the Q1 ("Did the app help you to solve your problem/achieve your goal?"), this study found that the maximum number of users (67.5%) have achieved their goals from App1 (Cervical Cancer Guide) followed by App2 (Figo Staging), App3 (Cervical Cancer Forum), and App4 (Cervical Cancer Tracker) 22.5%, 25.7%, 28.5%, and 67.5%.

Similarly, Figure 5 represents the user satisfaction regarding the UI design of the selected apps in response to the Q6
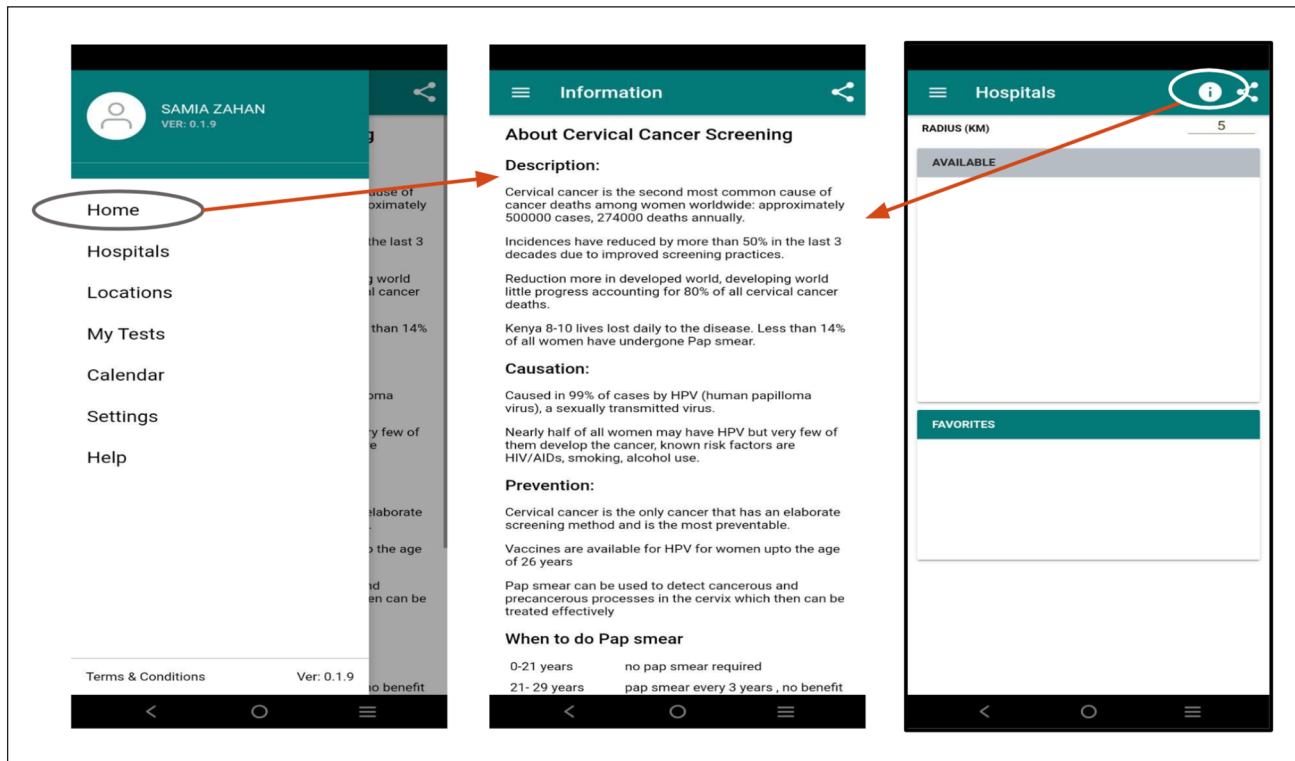
Figure 3. Screenshots of Cervical Cancer Tracker application.

(Which app has the most satisfying design (UI) as per you?) of G2 (UX). The result (see Table 7) showed that the UI design of App1 was comparatively better than the other three applications.

Again, in response to the Q7 (Which app is the least likable to you?) of G2 (UX) (see Table 7), the study showed that three out of 80 respondents voted for App1 as their least likable app; while maximum number (27.5%) of respondents voted for App3 and App4 as least likable app as shown in Figure 6.

In sum, the results of all 20 questions (see Tables 5–7) showed that App1 carries the good scores followed by App2, App3, and App4. The results thus indicated that App1 shows better performance followed by App2, App3, and App4 with respect to the six goals that includes user feedback on app's features, UX, comparative market values, design principles, error management, and overall feedback.

Comparing the outcomes

Some findings of GQM evaluation supported the findings through heuristic evaluation. In other words, a number of usability problems were revealed through both approaches while some were identical to each approach. For example, in Q3 (Which features did not work as expected?) of G1, a large number of users responded about the use of several unexpected icons and some functionalities such as diagnosis, staging, etc. Again, in response to Q13 (Is the system's icon

or representation relevant to you?) of G4 (Design Principles), 34 out of 80 respondents (43.75%) answered positively (yes) but the maximum of responses (56.25%) are negative (No) for App4 (Cervical Cancer Tracker) (see Table 5). These two questions (Q3 and Q13) subjectively match with H2 (Match between system and the real world). Similarly, in heuristic evaluation (see Table 4), the number of problems of App4 on H2 violation was three with an average severity of 3.33, which support 56.25% negative responses on Q13, which depicts similarity with the user responses of GQM study. Hence, these evaluations complement each other.

Again, Q16 (Does the system show easy search, limited number of buttons, control buttons, and clear terminology?) can be related to the H3 (User control and freedom). Hence, in GQM, users responded mostly positively for all the four apps. For App1, it is 98.5%. Again 97.5%, 98%, and 98.5% for App2, App3, and App4, respectively (see Table 5(c)), which indicated that the applications should not violate H3 much. In heuristic evaluation, we did not find any problem that violated H3 in App1, App2, and App3; as such the results of heuristic evaluation were similar to the user responses in terms of user's control and freedom. But in App4, there was a spike for H3 violation, involving the problem (No option to go back from current directory to root). This particular issue did not come up from the user response of Q16 but was detected as a violation in heuristic evaluation. Thus, here raised a deviation between GQM and heuristic evaluation. But this deviation was fixed by the responses of Q5 (What
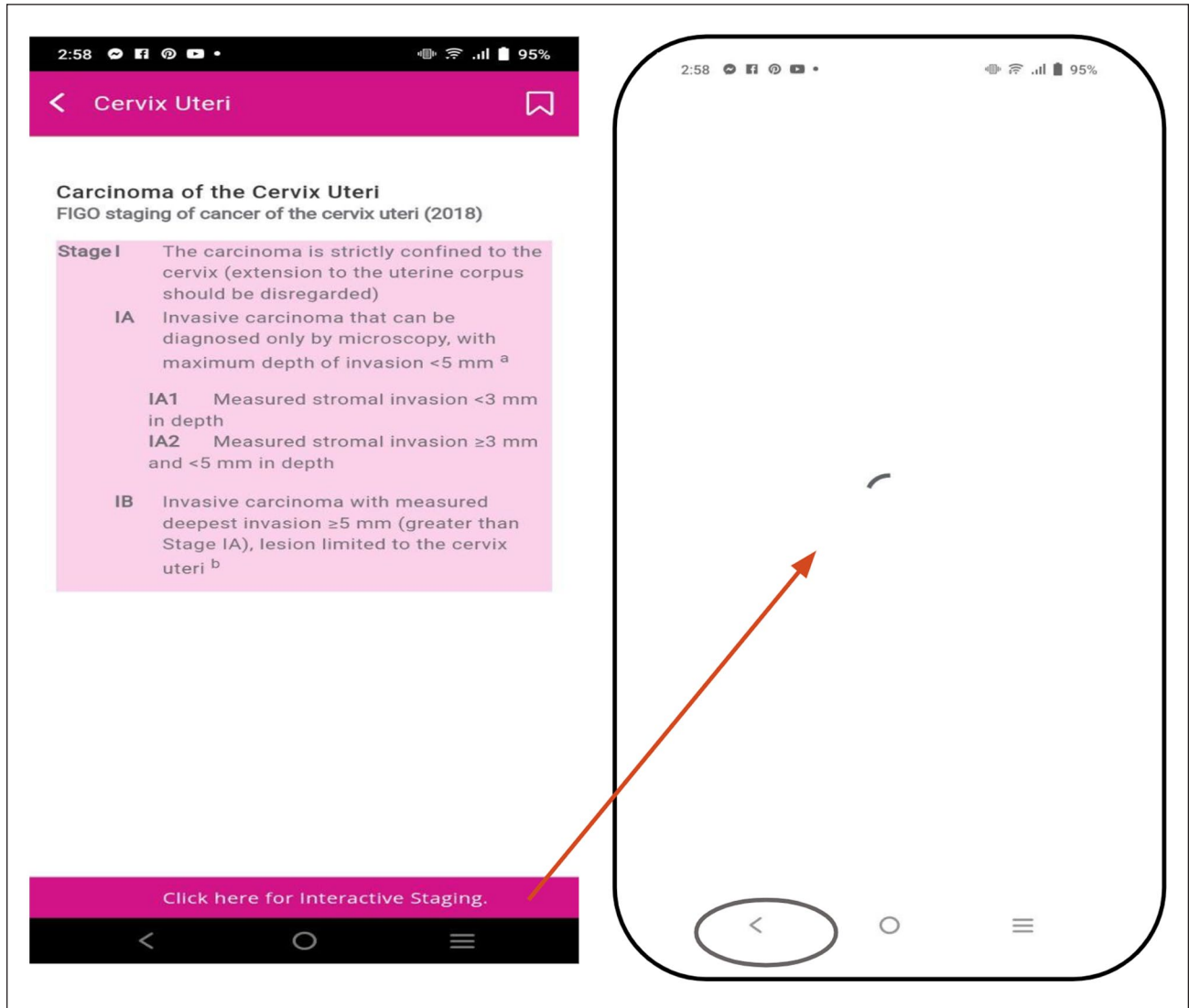
Figure 4. The interactive staging page of Figo Staging application.

Table 4. Summary results of heuristic evaluation.

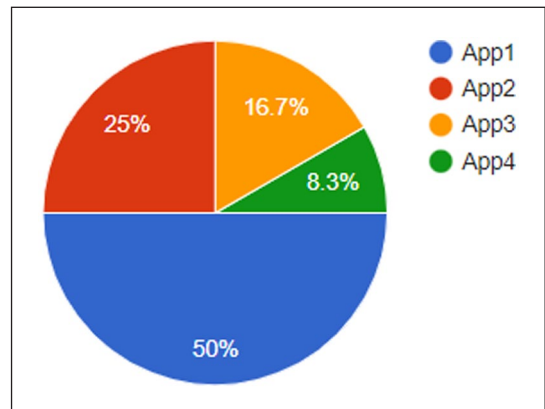| Heuristic	No. of problems				Average severity			
	App1	App2	App3	App4	App1	App2	App3	App4
H1	1	3	1	1	3	3.33	2	3
H2	0	1	1	3	0	3	4	3.33
H3	0	0	0	1	0	0	0	4
H4	3	0	1	2	3.33	0	4	4
H5	0	0	0	0	0	0	0	0
H6	0	0	0	0	0	0	0	0
H7	1	2	1	1	2	3	3	4
H8	4	2	4	3	3	2.5	3.25	3
H9	0	0	0	0	0	0	0	0
H10	0	0	0	1	0	0	0	4
Total	8	8	7	12	11.33	11.83	16.25	25.33

Figure 5. Users' satisfaction (in %) to the UI design of each application.

Table 5. Finings of GQM approach (dichotomous and semantic differential scale questions).

Goals	Questions	Metrics	Options/choices	Cervical Cancer Apps (values are in %)			
				App4	App3	App2	App1
G1: User feedback on app's features	Q1	Goal achievement	1 (Low)	15.3	22.3	12.5	0.0
			2	67.5	28.5	25.7	22.5
			3	8.7	18.4	19.0	4.3
			4	6.5	23.3	24.3	23.7
			5 (High)	2.0	7.5	18.5	49.5
	Q2	Attractiveness	1 (Less attractive)	43.7	10.9	11.2	2
			2	37.3	20.1	24.5	2.5
			3	2.0	12.7	19.3	4.5
			4	3.5	23.7	15.5	29.7
			5 (More attractive)	9.5	33.5	29.5	61.3
	Q3	Usefulness	Positive	2.3	25.5	47.9	98.9
			Negative	73.3	24.2	21.8	0.0
			N/A	24.4	50.3	30.3	1.1
	Q4	Emotional attachment	1 (Bad)	2.2	14.7	19.0	11.5
			2	4.1	10.5	9.3	5.3
			3	80.3	64.3	62.1	79.9
			4	7.9	5.5	4.4	2.0
			5 (Good)	5.5	5.0	6.2	1.3
G3: Comparative market values	Q9	App reachability	Friends	23.3	21.3	26	34.1
			Family	5.7	20.9	15	10.9
			Colleagues	5.7	2.7	25	13.5
			Social Media	65.3	55.1	34	41.5
	Q10	Competitiveness	1 (Worse)	69.3	32.5	1.4	1.1
			2	15.9	24.5	14.8	1.2
			3	6.6	15.7	21.5	2.3
			4	5.1	8.8	23.4	3.5
			5 (Better)	3.1	19.5	38.9	91.9
G4: Design principles	Q12	Visibility	Yes	23	63.8	72	95
			No	73	30.2	14	2
			N/A	4	6	14	3
	Q13	Relevance	Yes	43.75	92	91	96.5
			No	56.25	6.4	5.6	0.0
			N/A	0.0	1.6	3.4	3.5
	Q14	Consistency	1 (Low)	46.3	6.5	18	1.6
			2	33.5	10.7	12.7	3.5
			3	12.5	15.2	20.3	5.7
			4	7.7	24.3	20.5	10.2
			5 (High)	0.0	43.3	28.5	79
	Q15	Documentation	1 (Hard)	53.2	27.6	17.3	4.7
			2	39.3	40.1	10.2	5.3
			3	3.5	35.0	11.2	6.9
			4	1.5	25.0	25.9	15.1
			5 (Easy)	2.5	12.3	35.4	67.7
	Q16	Easy search, limited button/navigation link, intuitive terminology	Yes	98.5	98	97.5	98.5
			No	0.0	0.0	0.0	0.0
			N/A	1.5	2	2.5	1.5
G5: Error management	Q17	Error prevention	Yes	15.2	21.9	19	82
			No	84.8	58.0	40.5	0
			N/A	0.0	20.1	40.5	18

(Continued)

Table 5. (Continued)

Goals	Questions	Metrics	Options/choices	Cervical Cancer Apps (values are in %)			
				App4	App3	App2	App1
G6: Overall feedback	Q18	Quality	1	55.8	1.4	0.0	0.0
			2	32.7	3.9	6.0	0.0
			3	7.3	22.1	9.7	0.0
			4	4.2	25.6	32.4	2.8
			5	0.0	47.0	51.9	97.2
	Q19	Recommendation	1	55	0	0	0
			2	41	18	0	0
			3	4	20	0	0
			4	0	31	49	0
			5	0	31	51	100

Table 6. Finings of GQM approach (open-ended questions).

Goals	Questions	Metrics	Cervical Cancer Apps	Answers
G1: User feedback on app's features	Q5	Extension	App4	Treatment with images, address of hospitals, surgery, dashboard, customer feedback, personal account, e-medicine, live consultancy, life support, symptom analysis, back button
			App3	Doctor's info, donation, bonus, online services, adoption, precaution
			App2	track the record of camo therapy, synchronizing all doctor's info
			App1	Entertainment for patient
G2: UX	Q8	Negativity	App4	All functions
			App3	Notifications
			App2	Interface, speed
			App1	Treatment
G3: Comparative market values	Q11	Comparison	App4	Skin Vision, Facebook, Enriched website
			App3	Other websites
			App2	Cancer therapy advisor
			App1	N/A
G6: Overall feedback	Q20	Suggestion	App4	Need some more user friendly features, please focus on the interface and colors
			App3	Notifications feature should work properly
			App2	Highly expecting some design improvement
			App1	N/A

features would you like to add to the app?) of G1, where many users responded by seeking "Dashboard to roam between pages" and "Button for track back flexibility" features for App4 (see Table 6). Such responses also refer to the stated problem (No option to go back from current directory to root) that violated H3. This indicates that the responses of Q5 and Q16 together complement the Heuristic evaluation regarding the violation of H3.

Furthermore, in response to Q14 (Does the system look consistent (text, color) throughout the navigation?) of G4 (design principles), users mostly rated 1 (46.3%) or 2 (33.5%) for App4 and comparatively rated better for other Apps. Also, during the heuristic evaluation, several number of problems were observed like unnecessarily used different

Table 7. Finings of GQM approach (multiple choice question type close-ended questions).

Goals	Questions	Matric	Options/choices				
			App1	App2	App3	App4	N/A
G2: UX	Q6	Design (UI)	50	25	16.7	3.8	0
	Q7	Likability	3.8	18.8	27.5	27.5	22.4

font sizes, aesthetically bad design, poor color contrast, inconsistent behavior, etc., under App4 which supported this poor user rating for the violation of H8 (Aesthetic and minimalist design), H2 (Match between system and the real
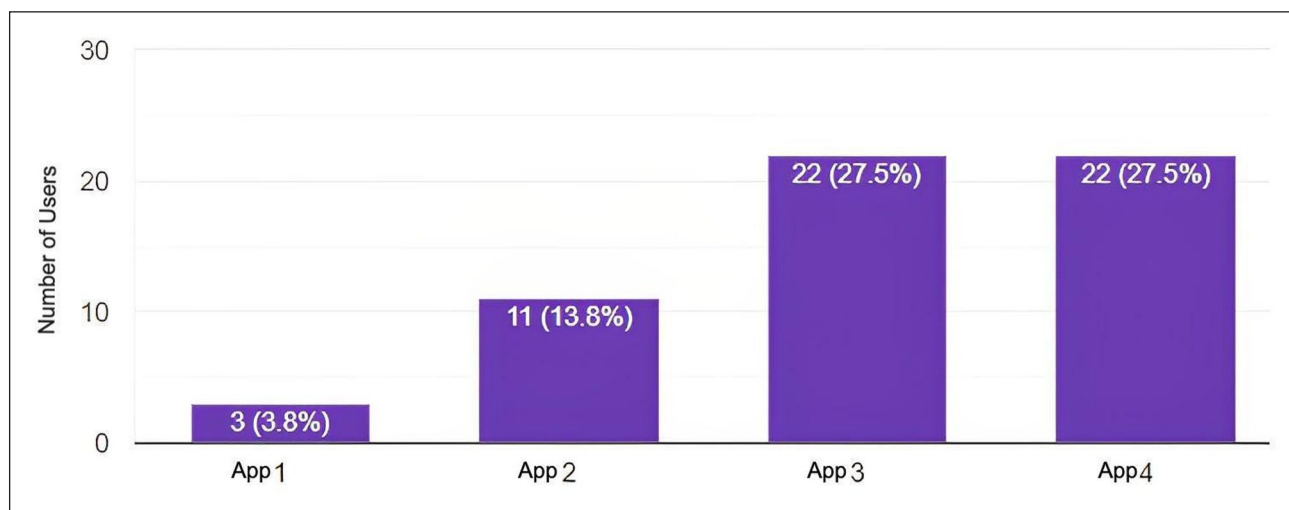
Figure 6. Users' likeability (in %) toward each application.

world), and H4 (Consistency and standards) in heuristic evaluation. A mapping between questions and heuristic violations was proposed based on the contextual similarities between them. For instance, H1 (Visibility of system status) that indicates the "visibility" concern is contextually similar to Q12 (Is there a clearly identified link to the home page?) which was asked for measuring the metric named "visibility" (see Table 2). Again, the H3 (Error prevention) and the Q17 (Is the application alerting you for giving any wrong interaction/input?) which was asked to measure the metric "error prevention" (see Table 2) focusing on the similar context. Thus, the questions that are complementing the heuristics are mapped as presented in Table 8.

Apart from the complementing questions, some other questions cover the additional aspects for the evaluation of the apps. For example, how effective and informative an application is, this cannot be measured from Neilson's Heuristic evaluation. Also the importance of these apps with respect to usefulness and how much users are emotionally attached or dependent on these apps, cannot be evaluated by heuristic violation finding. Enhancement scopes of the apps, flourishing edges, and overall performance measures are also very important aspects for a wholesome evaluation of any application. These cannot be obtained by Neilson's HE. Therefore, the GQM questionnaires include some questions to assess these issues which augments the heuristic evaluation. For example, in the GQM approach, the effectiveness of the selected apps was measured through the question Q1 (Did the app help you to solve your problem/achieve your goal?) of G1 (User Feedback on App's Features). The study found that 22.5%, 25.7%, 28.5%, and 67.5% of respondents rated 2 out of 5 for app1, app2, app3, and app4, respectively (see Table 5). Although we obtained that App1 is better rated than other apps through heuristic evaluation, this below

Table 8. GQM-based questions that complements Neilson's heuristics.

Heuristics	Questions
H1	Q12
H2	Q3, Q13
H3	Q5, Q17
H4	Q3, Q6, Q15
H5	Q3, Q17
H6	Q3
H7	Q2, Q3, Q8, Q9
H8	Q2, Q3, Q7
H9	Q3, Q16
H10	Q3, Q16

Table 9. GQM-based questions that augmented Neilson's heuristics.

Questions	Measures
Q1	Effectiveness
Q4, Q11	Importance of App with respect to usefulness and emotion
Q5, Q18	Enhancement of the app
Q10	Flourishment of the app
Q13	Overall performance of the app

average rating on Q1 from user end indicates that these apps are not very effective yet for a considerable percentage of users. These may be performing better in respect to heuristic evaluation aspects but actually cannot serve the purposes as expected. As such, GQM analysis augments heuristic evaluation. Table 9 shows the mapping between augmented questions and measuring aspects.

Discussion

Main findings

Evaluating the usability of cervical cancer applications was the primary aim of this study. Among the selected applications, the "Cervical Cancer Guide" (App1) is performing comparatively better, while the "Cervical Cancer Tracker" (App4) showed the lowest value in the performance ranking. Both the approaches (GQM and HE) derive to almost similar outcomes (see Figure 5 for GQM approach outcome and Table 4 for the average severity in heuristic approach). The study also showed that none of the approach individually able to offer complete evaluation of an application. In GQM, some questions were added that do not complement any heuristic violation directly or indirectly. The responses of these questions showed that none of these apps are actually fulfilling users' expectations and demands completely. For example, in response to Q4 (How would you feel if you can no longer use the app?) of G1, most of the respondents responded neutral, that is rating 3 (see Table 5). Which indicates, these applications are still not benefiting people the way they are expected to. Another finding is, it is not preferable to the users to roam between apps for different purposes like, cervical cancer basic, cervical cancer symptoms, screening, staging, treatment, etc. Therefore, an application uniting all the purposes together is more convenient and acceptable to the users. Again, in response to Q5 (What features would you like to add to the app?), respondents proposed different features which indicated the requirement of developing an application that includes the all necessary features.

The research domain, cervical cancer, is an important sensitive area of development in the recent era. There are huge scopes to develop more informative, effective, efficient, and user friendly apps of this domain in future. The results imply that the usability of applications for cervical cancer is unsatisfactory for both expert-based and user-based evaluation. In this research, the issues and data that are found from the four cervical cancer applications can be utilized by designers to increase the usability of their applications. This study demonstrated that the number of usability issues discovered by expert-based HE and user-based "GQM" methodology through questionnaires did not differ considerably. The user-based approach, however, performs noticeably better in finding usability issues that imply the systems learnability, whereas expert-based heuristic evaluation performs better in identifying issues that cause user annoyance.

Comparison with prior work

A comparative discussion with the related prior works are presented here. First, in some research, authors developed and proposed a mobile application focusing on specific themes such as screening, awareness,[13,15] while we have not designed or developed any app but evaluated four existing apps focusing on two different themes. In our research, we have carried out usability evaluation and highlighted the issues to improve system design. Thus, we tried to consider multiple related themes and come up with a convenient and user friendly application buildup idea through comparative analysis rather than offering a concrete prototype. Thus, this study can be benefitted to design user-intuitive interfaces for developing cervical-cancer-related applications focusing on multiple themes.

Second, earlier studies emphasized the importance of mobile apps in terms of knowledge gaining on cervical cancer domain. As an instance, Lee et al.[14] revealed that mobile apps are effective tools to enhance knowledge regarding cervical cancer. In our research, we have chosen cervical-cancer-related mobile applications and evaluated those to help users get a convenient and user friendly application to gain diverse knowledge on cervical cancer, since the selected four apps share information on guidelines, forum, tracking, and staging of cervical cancer.

Finally, only a few studies conducted the evaluation while some highlighted the necessity of evaluating the mobile apps.[4] In our studies, we have explicitly focused on evaluating the developed application from a usability perspective to highlight not only the usability problems rather its importance in designing and developing the cervical cancer-related applications.

Limitations

This research has few limitations as well. The number of respondents was not sufficient to pursue the GQM study. Power calculation was not considered to analyze the study data. Again, no pilot testing has been conducted for validating the questions before the final stage though the questionnaires were prepared based on the GQM structure.[30] Before distribution, the questions were verified and refined based on the feedback taken from peers and experts. Again, only four apps belonging to two themes were evaluated in this study.

Future work

In future research, instead of two themes, more applications focusing on other themes such as consultancy, treatment, etc. could be chosen to make the results more generalized. Likewise, several apps of each theme may be chosen. The data analysis may be conducted for individual thematic apps separately and come up with the best possible design solution of each thematic app. And finally proposing an application which will combine all the themes and serve the user with utmost convenience. For the evaluation, some other approaches like laboratory based usability testing, cognitive walkthrough, etc. may be considered. Similarly, more experts and respondents from diverse backgrounds will help to find a more generalized and better result.

Conclusions

Today's medical research is not entirely in line with all the new developments and breakthroughs. In terms of the effort that researchers give[39–41]; whether it be in preclinical research or trials, so much of it is still manual. However, this research will be beneficial for future researchers on cervical cancer with respect to design and develop a usable and intelligent system. Again, since mobile health application is an integral part of treatment,[42–44] mobile application developed for the treatment of cancer patients may bring the doctors, nurses, and patients at a single platform at the same time, which will be a blessing for all of them. As such, such applications should be easily accessible and usable to the all focused users (doctors, nurses, and patients) so that they can use the applications without any hassle but effectively, efficiently and with satisfaction.

Author contributions

- **Md Abdur Razzak:** Conceptualization; Methodology; Investigation; Writing—review & editing.
- **Muhammad Nazrul Islam:** Conceptualization; Formal analysis; Investigation; Methodology; Supervision; Writing—review & editing.
- **Tanjila Broti:** Investigation; Methodology; Writing—original draft; Writing—review & editing.
- **Elma Shumsun Kamal:** Data curation; Investigation; Writing—review & editing.
- **Samia Zahan:** Methodology; Writing—review & editing.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethics approval and consent to participate

We confirm that ethical approval was applied for conducting this research. The target of the study was to collect survey responses from human participants on the subjective usability assessment of the selected mobile applications. No human data, human tissue, or any clinical data were collected for this study. Therefore, the ethical committee headed by the Research & Development Wing of Military Institute of Science and Technology (MIST) decided that it is not required to have formal approval. We also declare that we have taken written consent from survey respondents to participate in the evaluation study.

Informed consent

Written informed consent was obtained from all subjects before the study.

Trial registration

Not applicable.

ORCID iD

Muhammad Nazrul Islam [iD] https://orcid.org/0000-0002-7189-4879

References

1. Hu L, Horning MP, Banik D, et al. Deep learning-based image evaluation for cervical precancer screening with a smartphone targeting low resource settings–engineering approach. In *2020 42nd Annual international conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, July 20-24, 2020, pp. 1944–1949, Montreal, Canada: IEEE.
2. Rao H-L, Chen J-W, Li M, et al. Increased intratumoral neutrophil in colorectal carcinomas correlates closely with malignant phenotype and predicts patients' adverse prognosis. *PLoS One* 2012; 7(1): e30806.
3. Cohen PA, Jhingran A, Oaknin A, et al. Cervical cancer. *The Lancet* 2019; 393(10167): 169–182.
4. Ruiz-Lo'pez T, Sen S, Jakobsen E, et al. Fighthpv: design and evaluation of a mobile game to raise awareness about human papillomavirus and nudge people to take action against cervical cancer. *JMIR Serious Games* 2019; 7(2): e8540.
5. Endale H, Mulugeta T and Habte T. The socioeconomic impact of cervical cancer on patients in Ethiopia: evidence from Tikur Anbessa Specialized Hospital. *Cancer Manage Res* 2022; 14: 1615.
6. Ormel I, Onu CC, Magalhaes M, et al. Using a mobile app–based video recommender system of patient narratives to prepare women for breast cancer surgery: development and usability study informed by qualitative data. *JMIR Form Res* 2021; 5(6): e22970.
7. Ginossar T, Shah SFA, West AJ, et al. Content, usability, and utilization of plain language in breast cancer mobile phone apps: a systematic analysis. *JMIR mHealth and uHealth* 2017; 5(3): e7073.
8. Vergani L, Marton G, Pizzoli SFM, et al. Training cognitive functions using mobile apps in breast cancer patients: systematic review. *JMIR mHealth and uHealth* 2019; 7(3): e10855.
9. Quercia K, Tran PL, Jinoro J, et al. A mobile health data collection system for remote areas to monitor women participating in a cervical cancer screening campaign. *Telemedicine and e-Health* 2018; 24(4): 277–282.

10. Champin D, Ram'ırez-Soto MC and Vargas-Herrera J. Use of smartphones for the detection of uterine cervical cancer: a systematic review. *Cancers* 2021; 13(23): 6047.

11. Moodley J, Constant D, Botha MH, et al. Exploring the feasibility of using mobile phones to improve the management of clients with cervical cancer precursor lesions. *BMC Women's Health* 2019; 19(1): 1–10.

12. Shanthi P, Hareesha K and Kudva R. Automated detection and classification of cervical cancer using pap smear microscopic images: a comprehensive review and future perspectives. *Engineered Science* 2022; 19: 20–41.

13. Muljo HH, Perbangsa AS, Paedamean B, et al. Mobile learning for early detection cancer. *Int J Interact Mobile Technol* 2018; 12(2): 39 –53.

14. Lee HY, Koopmeiners JS, Rhee TG, et al. Mobile phone text messaging intervention for cervical cancer screening: changes in knowledge and behavior pre-post intervention. *J Med Internet Res* 2014; 16(8): e3576.

15. Stocks J, Choi Y, Ibrahim S, et al. Iterative development of a mobile phone app to support community health volunteers during cervical cancer screening in western Kenya: qualitative study. *JMIR Formative Res* 2022; 6(2): e27501.

16. Bhochhibhoya S, Dobbs PD and Maness SB. Interventions using mhealth strategies to improve screening rates of cervical cancer: a scoping review. *Prevent Med* 2021; 143: 106387.

17. Kanjak J, Likitdee N, Kietpeerakool C, et al. Evaluation of mobile health applications for cervical cancer in the digital marketplace. *Obstet Gynecol Sci* 2022; 65(3): 244–255.

18. Kazemi S, Zarei F, Heidarnia A, et al., Improve the cervical cancer prevention behaviors through mobile-based educational intervention based on I-CHANGE model: study protocol for a randomized controlled trial. *Trials* 2022; 23(1): 805.

19. Islam MN, Karim M, Inan TT, et al. Investigating usability of mobile health applications in bangladesh. *BMC Med Inform Decis Mak* 2020; 20(1): 1–13.

20. Yang Y, Chen H, Qazi H, et al. Intervention and evaluation of mobile health technologies in management of patients undergoing chronic dialysis: scoping review. *JMIR mHealth and uHealth* 2020; 8(4): e15549.

21. Omar KS, Mondal P, Khan NS, et al. A machine learning approach to predict autism spectrum disorder. In *2019 International conference on electrical, computer and communication engineering (ECCE)*, February 7-9, 2019, pp. 1–6, Cox's Bazar, Bangladesh: IEEE.

22. Khan NS, Muaz MH, Kabir A, et al. Diabetes predicting mhealth applications using machine learning. In *2017 IEEE international WIE conference on electrical and computer engineering (WIECON-ECE)*, December 18-19, 2017, pp. 237–240, Uttarkhand, India: IEEE.

23. Khan NI, Mahmud T, Islam MN, et al. Prediction of cesarean childbirth using ensemble machine learning methods. *In Proceedings of the 22nd international conference on information integration and web-based applications & services*, 30 November–02 December, 2020, pp. 331–339, Chiang Mai, Thailand

24. Tseng C-J, Lu C-J, Chang C-C, et al. Application of machine learning to predict the recurrence-proneness for cervical cancer. *Neural Comput Appl* 2014; 24(6): 1311–1316.

25. Islam MN and Tétard F. Integrating semiotics perception in usability testing to improve usability evaluation. In *Cases on usability engineering: design and development of digital products*, IGI Global, 2013, pp. 145–169, PA, USA.

26. Islam MN, Khan SR, Islam NN, et al. A mobile application for mental health care during covid-19 pandemic: development and usability evaluation with system usability scale. In: *Computational Intelligence in Information Systems: Proceedings of the Computational Intelligence in Information Systems Conference (CIIS 2020), January 25-27, 2021, pp. 33–42, Darussalam, Brunei: Springer International Publishing*.

27. Rahman MM, Sarker A, Khan IB, et al. Assessing the usability of ridesharing mobile applications in Bangladesh: an empirical study. In *2020 61st International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*, October 15-16, 2020, pp. 1–6. Riga, Latvia: IEEE.

28. Monkman H and Kushniruk A. A health literacy and usability heuristic evaluation of a mobile consumer health application. In *Porceedings of the 14th World Congress on Medical and Health Informatics (MEDINFO 2013) MEDINFO 2013*, August 2013, pp. 724–728, Copenhagen, Denmark: IOS Press.

29. Bhowmik M, Ashraf F, Fatema T., et al. Evaluating usability of mobile financial applications used in Bangladesh. In *Advances in Design and Digital Communication III: proceedings of the 6th International Conference on Design and Digital Communication, Digicom 2022*, November 3–5, 2022, Barcelos, Portugal, pp. 161–176. Cham: Springer Nature Switzerland.

30. Hussain A and Ferneley E. Usability metric for mobile application: a goal question metric (GQM) approach. In: *Proceedings of the 10th international conference on information integration and Web-Based Applications & Services*, 2008, pp. 567–570.

31. Caldiera VRBG and Rombach HD. The goal question metric approach. In *Encyclopedia of Software Engineering*, 1994, pp. 528–532.

32. Ssemugabi S and De Villiers M. Effectiveness of heuristic evaluation in usability evaluation of e-learning applications in higher education. *S Afr Comput J* 2010; 2010(45): 26–39.

33. Kumar BA, Goundar MS and Chand SS. A framework for heuristic evaluation of mobile learning applications. *Educ Inform Technol* 2020; 25(4): 3189–3204.

34. Gonzalez-Holland E, Whitmer D, Moralez L, et al. Examination of the use of nielsen's 10 usability heuristics & outlooks for the future. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2017, vol. 61, no. 1, pp. 1472–1475, Los Angeles, CA: SAGE.

35. Bokingkito PB Jr and Caparida LT. Usability evaluation of a real-time water quality monitoring mobile application. *Proc Comput Sci* 2022; 197: 642–649.

36. Hussain A and Ferneley E. Usability metric for mobile application: a goal question metric (GQM) approach. In *Proceedings of the 10th international conference on information integration and web-based applications & services*, November 24-26, 2008, pp. 567–570, Linz, Austria.

37. Hussain A and Kutar M. Usability evaluation of SatNav application on mobile phone using mGQM. *Int J Comput Inform Syst Ind Manage Appl* 2012; 4(2012): 92–100.

38. Santos BS, Silva S and Dias P. Heuristic evaluation in visualization: an empirical study: Position paper. In: *2018 IEEE evaluation and beyond-methodological approaches for*

visualization (BELIV), October 21, 2018, pp. 78–85, Berlin, Germany: IEEE.

39. Inan TT, Samia MBR, Tulin IT, et al. A decision support model to predict ICU readmission through data mining approach. In: *Proceeding of the Pacific Asia conference on information systems, AIS*, Article No. 218, June 26-30, 2018, Yokohama, Japan.

40. Omar KS, Anjum A, Oannahary T, et al. An intelligent assistive tool for Alzheimer's patient. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, May 03-05, 2019, pp. 1–6, Dhaka, Bangladesh: IEEE.

41. Suha SA and Islam MN. An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image. *Sci Rep* 2022; 12(1): 17123.

42. Suha SA, Islam MN, Akter S, et al. Assessing usability of mobile applications developed for autistic users through heuristic and semiotic evaluation. In: *Proceedings of International Joint Conference on Advances in Computational Intelligence: IJCACI 2021*, May 2022, pp. 25–39. Singapore: Springer Nature Singapore.

43. Hasan U, Nazrul Islam M, Tajmim Anuva S, et al. User-centred design-based privacy and security framework for developing mobile health applications. In *Proceedings of international joint conference on advances in computational intelligence: IJCACI 2020*, 2021,November 20-21, 2020, pp. 203–216, Springer Singapore.

44. Islam MN, Islam I, Munim KM, et al. A review on the mobile applications developed for COVID-19: an exploratory analysis. *IEEE Access* 2020; 8: 145601–145610.