Usability Assessment of the Missouri Cancer Registry's Published Interactive Mapping Reports: Round Two

Awatef A. Ben Ramadan, MD, MPH[1,2,3,4], Jeannette Jackson-Thompson, MSPH, PhD[2,3,4], Chester L. Schmaltz, PhD[2,3]

[1]Department of Mathematics, Science, and Informatics, College of Professional Advancement, Mercer University

[2]Missouri Cancer Registry and Research Center (MCR-ARC), University of Missouri-Columbia (MU),

[3]Department of Health Management and Informatics (HMI), School of Medicine, University of Missouri-Columbia (MU)

[4]MU Informatics Institute (MUII), University of Missouri-Columbia (MU)

Abstract

Background: Health-related data's users have trouble understanding and interpreting combined statistical and mapping information. This is the second round of a usability study conducted after we modified and simplified our tested maps based on the first round's results.

Objective: To explore if the tested maps' usability improved by modifying the maps according to the first round's results

Methods: We recruited 13 cancer professionals from National American Central Cancer registries (NACCR) 2016 conference. The study involved three phases per participant: A pretest questionnaire, the multi-task usability test, and the System Usability Scale (SUS). Software was used to record the computer screen during the trial and the users' spoken comments. We measured several qualitative and quantitative usability metrics. The study's data was analyzed using spreadsheet software.

Results: In the current study, unlike the previous round, there was no significant statistical relationship between the subjects' performance on the study test and the experience in GIS tools (P = .17 previously was .03). Three out of the four (75%) of our subjects with a bachelor's degree or less accomplished the given tasks effectively and efficiently. This study developed a comparable satisfaction results to the first round study, despite that the previous round's participants were highly educated and more experienced with GIS.

Conclusion: By considering the round one's results and by updating our maps, we made the tested maps simpler to be used by subjects who have little experience in using GIS technology, and have little spatial and statistical knowledge.

Keywords: Geographic Information Systems, Interactive Maps, Missouri Cancer Registry, NAACCR, Usability

Abbreviations: GIS: Geographic Information System

HMI: Health Management and Informatics

IRB: Institutional Review Board

MCR-ARC: Missouri Cancer Registry and Research Center

MPH: Master of Public Health

NAACCR: North America Association of Central Cancer Registries

ORE: Overall Relative Efficiency

SUS: System Usability Scale

TBE: Time-Based Efficiency

TCR: Task Completion Rate

Correspondence: 3001 Mercer University Drive., AACC Building

Atlanta, Georgia, 30341

Benramadan_aa@mercer.edu

Office phone: (678) 547-6240

DOI: 10.5210/ojphi.v11i2.9483

Introduction

There are enormous Geographic Information System (GIS) technologies that have been created and designed to visualize different kinds of health and health-related data. These tools should be designed and modified to meet the perceptions and needs of these technology's possible users [1]. Software developers and designers, as well as the GIS technology innovators should concentrate on how to make this technology effective and efficient for the targeted users [2,3].

Public health professionals have started using specific advanced software to examine and illustrate population-based databases. Illustrating and visualizing this kind of information becomes essential and important to make a measurable impact on the public health problems in any community. This type of technology has been influential on public health research, as well as on development of new effective public health policy. Therefore, we have to be sure that the population-based databases are would be analyzed intelligently and examined appropriately to yield reliable outcomes, and that the results do not mislead the targeted audiences [4].

The previous usability literature pointed out that most of any new numerical technology users face difficulties in interpreting the combined and multiple sources data [5-7]. The same difficulties have challenged new GIS tools because of the combined geographical and statistical data of these tools. The scientifically proved interpretation was: Insufficient knowledge and inadequate hand-on experience on using GIS tools, ignorance and a decline to practice this technology between its prospective users, and because there were usability issues and complexity perception among the possible users towards this kind of technology [8].

As previous scientific research revealed, regular and interactive mapping reports can produce knowledge, create evidence, and augment strategies [9]. Therefore, every interactive mapping report should carry a clear aim and convey a definite message to the targeted audiences [10]. These interactive maps must include citations and details of used information resources and the detailed methodology which followed the production of visualized results. The GIS technology related literature revealed that the mapping reports should undergo strong scrutiny and evaluation, using representative samples of the users, to assess the usability and make this technology fit the users' needs and preferences [4]

The collaboration between public health experts and the other specialists of the same research and practice interests has been proven scientifically and this relation's positive impact on the public health problems and disparities has been confirmed [11]. Using highly advanced and sophisticated GIS tools will help public health experts to plan and create cost-effective health-related strategies and policies [11]. During the third millennium, the health related maps have changed from being static to be interactive [12]. The scientific literature discovered that the GIS technology users prefer dynamic interactive maps [4]. The same literature concluded that the atlas creators need to take in their account the potential users' preferences and perceptions. The map developers should consider the users' needs starting from the planning and designing processes, and ending by evaluating the maps before and after releasing them to the actual users [13].

Many cancer registries have established interactively visualizing their data' outcomes but small number of them are evaluating the usability of their maps [14-16]. By conducting this two-round study we seek to fill this breach and give a standard model in evaluating cancer registries' maps and help in tailoring these tools according to the potential users' insight and favorites.

In the first round of the study, the investigators conducted a usability testing trial to assess the usability of the two published Missouri Cancer Registry and Research Center (MCR-ARC)'s interactive reports. The first round's participants were faculty and staff of Master of Public Health program (MPH) and the Department of Health Management and Informatics (HMI) at University of Missouri-Columbia [17].

In the current study, we conducted a second round of the same first round's methodology to measure effectiveness and efficiency of the published InstantAtlas reports of MCR-ARC after we modified them according to the first round results [17]. In this round, we selected a convenience sample of cancer professionals who attended the North America Association of Central Cancer Registries (NAACCR) conference-June, 2016. The second round of study also aims to evaluate if and for how extent the users' action will be influenced by the users demographic information, experience, education level, and the work type.

Methods

Study Design

The investigators selected a mixed methodology tactic. The tested mapping reports are already published on the MCR-ARC (see Multimedia Appendices 1 and 2) [17-19]. The trial elements were the same Round One's components: A pretest questionnaire, the multi-task usability test, and the System Usability Scale (SUS) per every participant respectively [17,20].

The pretest questionnaire

The questionnaire involved inquiries on every subject's age, race, work type, education level, total experience in public health field, and experience in GIS tools use (see Multimedia Appendix 3). This step was followed by the multi-task test.

Multi-Task usability test

This stage included ten independently ranked tasks, which cover most of the maps' functionality. These tasks were handled by the study subjects individually to test the usability of the tested InstantAtlas reports. The Multi-Task scenario was constructed by the study's investigators grounded on the anticipated functionality of the tested maps. This phase aimed to precisely estimate the efficiency and effectiveness in terms of task completion success rate and task completion time [21]. All the tasks were in the same classification and context for all subjects (see Multimedia Appendix 4).

The System Usability Scale (SUS)

This phase was composed of a manufacturing and simple ten-item scale to assess the subjects' independent assessment of the experienced mapping reports' usability. This phase was immediately conducted after the Multi-Task scenario phase. The SUS score range is between 0 and 100. Sixty-eight or above has been considered as satisfactory based on previous usability literature and upper score up to 100 is the optimum to finest score [20]. Scores above 68 points are acceptable according to usability scholars and higher scores represent the optimal to best score [20].

Participants

The study's protocol was accepted by the Health Sciences Institutional Review Board (IRB) of the University of Missouri-Columbia. The current round's primary investigator attended the NAACCR-2016 conference and she convinced group of attendees to participate in the study and conduct the study's trial. The convenience sampling method was selected to collect the current study's participants. The investigators conducted the study's experiment on the first reacted thirteen participants. As the usability scholars confirmed, five was the smallest number of subjects to run a fruitful usability study; a five subjects study enables revealing from 55 to 100% of the usability issues of any experienced material [22]. The investigators increased the number to thirteen subjects to expose as many as probable of the usability issues of our refined published mapping reports [23].

Study Procedure

Each subject tried ten tasks in a secure place for an average of 30 minutes per participant. A specific computer laptop was used to conduct the study. The researchers installed a Microsoft Windows-7 software, Windows Media Player, to audio-video record the laptop's screen while the subjects performing the experiment. Task on time and task completion success rates were analyzed manually based on the recordings.

The subsequent usability metrics were also measured:

Performance Metrics

Some metrics were measured to evaluate the effectiveness and the efficiency of the tested maps and to explore the maps' usability issues. The critical error is when the participant required assistance from the test viewer to finish a task [21]. The investigators measured the following metrics:

Effectiveness, Task Completion Rate (TCR): TCR is a measure of tasks that were completed without critical errors, and the outputs of the task were correct [25]. TCR equals to the effectiveness of the tested maps and was represented in two different means: by participant and by task.

TCR per participant: The percentage of tasks that were successfully completed by a participant [24].

$$\text{TCR per participant} = \frac{\text{Number of tasks completed successfully}}{\text{Total number of tasks undertaken}} \cdot 100\%$$

TCR per task: The percentage of participants who successfully completed a given task [24].

$$\text{TCR per task} = \frac{\text{Number of participants who completed successfully}}{\text{Total number of participants}} \cdot 100\%$$

Efficiency: The efficiency is defined according to the ISO-9241 as: "Resources spent by user in order to ensure accurate and complete achievement of the goals" [25].

The investigators calculated the efficiency and the productivity of the tested mapping reports using the following couple of formulas

Time Based Efficiency (TBE) [24]

$$\bar{P}_{t,j} = \frac{1}{N}\sum_{i=1}^{N}\frac{n_{ij}}{t_{ij}}$$

$N =$ Total number of tasks

$n_{ij} =$ Result of task i by user j

$$= \begin{cases} 1 & \text{if the user successfully completes the task} \\ 0 & \text{otherwise} \end{cases}$$

$t_{ij} =$ The time spent by user j on task i.

Overall Relative Efficiency (ORE) [24]

$$\bar{P}_{j} = \frac{\sum_{i=1}^{N} n_{ij} t_{ij}}{\sum_{i=1}^{N} t_{ij}} \cdot 100$$

User Satisfaction

SUS scale was used to assess the satisfaction per study subject [20]. See the detailed SUS mechanism under the study design section.

Factors affecting the participants' performance

The current study researchers measured some factors those they expected might impact the participants' performance and satisfaction during the trial [26]. Those factors were: the participants' education level, work type, experience in healthcare field, and previous experience with mapping reports and GIS tools. The investigators chose different statistical measures, as needed, to assess the chosen factors' relationships (Wilcoxon-Mann-Whitney test, Pearson correlation test, and/or simple regression test) [27]. The intended sample size of this study was small since we primarily wished to uncover major usability problems; post-hoc power calculations for simple linear regression with the observed sample data indicates that the power for testing the relationships between the participants' factors and the TCR or SUS ranged between 7% and 32% [28]. We used a type I error rate (α) of 0.05 for the hypothesis tests conducted in this project.

Results

1. Participant demographics

The study researchers interviewed thirteen cancer health professionals, four males and nine females. Their ages ranged from 29-56 years old (Mean=40.85 years old, Median=40 years old). All of the thirteen participants were cancer professionals who attended NAACCR-2016 conference. The race of our participants was as following: seven Whites, four Blacks, one Asian, and one Native Hawaiian or other Pacific Islander. Three of the subjects hold PhD degrees, six hold master degrees, one bachelor holder, one associate degree holder, and only one got some college level education. The associate degree and the some college holders were working as certified cancer registrars in two separate central cancer registries. At the time of data collection, the PhD holder subjects' work roles were: a cancer researcher, an epidemiologist, and a director and associate professor. The master degree holders' work roles were: a program director and epidemiologist, a statistical consultant, a research coordinator, a software engineer, a cancer surveillance and epidemiologist, and an epidemiologist. The bachelor degree holders were functioning as: a research analyst and public health epidemiologist, and an abstractor and quality control manager. The associate degree holder was functioning as a certified cancer registrar and the subject who holds some college education was working as a data manager in a central cancer registry. The total experience in public health for the all subjects ranged between 2 years to 19 years (Average= 10.38 years, Median= 10 years). The subjects' total experience in using GIS tools in work was between null experience to 10 years (Mean= 2 years, Median= 0 years).

2. Reports' effectiveness and efficiency

The mapping reports' effectiveness

Effectiveness per Participant

One of the participants could not complete three successive tasks #4, 5, & 6. The subject held a bachelor degree and she did not have any previous experience in using GIS tools and interactive mapping reports. Another subject could not get the assigned results for two successive tasks #4 & 5. This subject holds a PhD degree and have had an experience in public health for 15 years, an experience in using GIS tools at work, and work as a director at one of the central cancer registries. Two of the subjects could not complete two non-successive tasks: one of these subjects holds master's degree and could not finish the tasks #5 & 8 successfully and another subject holds just some college education missed the tasks #4 & 8. The master's degree subject had experience in public health field for 15 years and five years of using GIS tools and software, and the some college degree holder had experience of 19 years in public health and no previous experience in using or reading GIS tools. Four of our subjects could not accomplish just one task and the missed tasks were #5, 6, &9. Two of these participants were PhD holders, one master degree holder, and a bachelor degree holder. The two PhD holders had experience of 8 and 10 years in public health field and experience of 4 and 5 years of using GIS tools in their daily work. The master degree holder had 17 years of public health experience and no previous experience in GIS tools. The bachelor degree holder had experience in public health field of 10 years with null experience in using GIS tools.

Only four of our 13 participants (30.76%) achieved the TCR of 100%. Three of these participants were master's degree holders and one of them held an associate degree in science. These four subjects hold experience in public health field between two to ten years, and all of them carried no previous experience in using GIS tools and interactive mapping reports. Twelve of our 13 participants (92.30%) were able to accomplish the tasks with TCR of 79% or above. The only one subject who got TCR of <79% was a PhD holder and a director of a central cancer registry with 12 years public health experience and null GIS experience.

The results were ranged from 70% to 100% as Figure (1) shows.

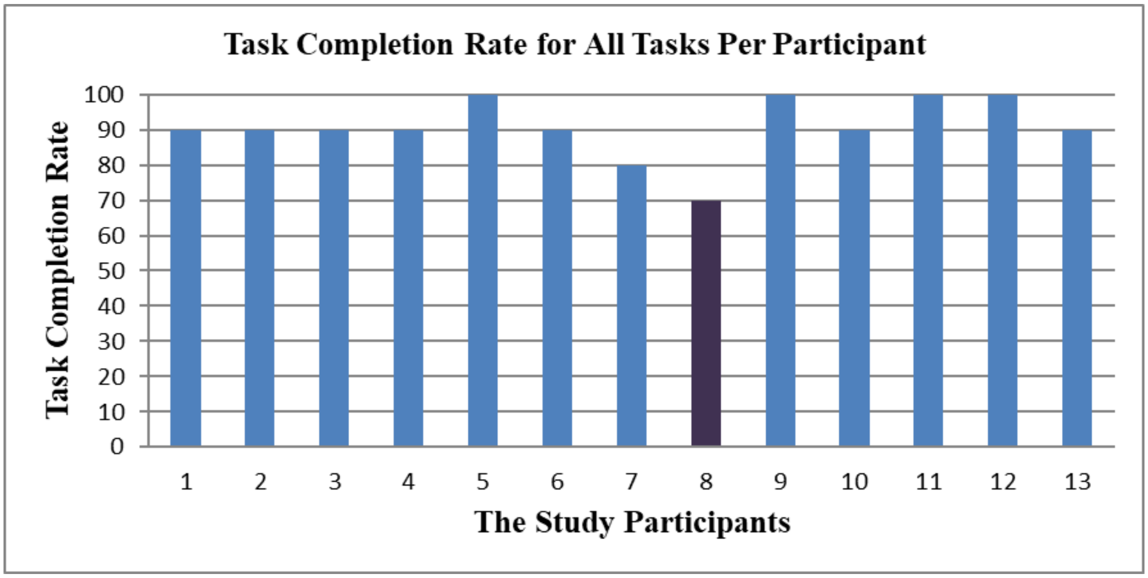


Figure (1). Task Completion Rate for All Tasks per Participant. Blue bars indicates participants who finished the trial with > 78% TCR; Purple bar indicates a participant who finished the trial with <78% TCR.

Effectiveness per Single Task

The above task completion formula were also used to calculate the task completion rate per task of the study's ten tasks. The results are presented in Figure (2)

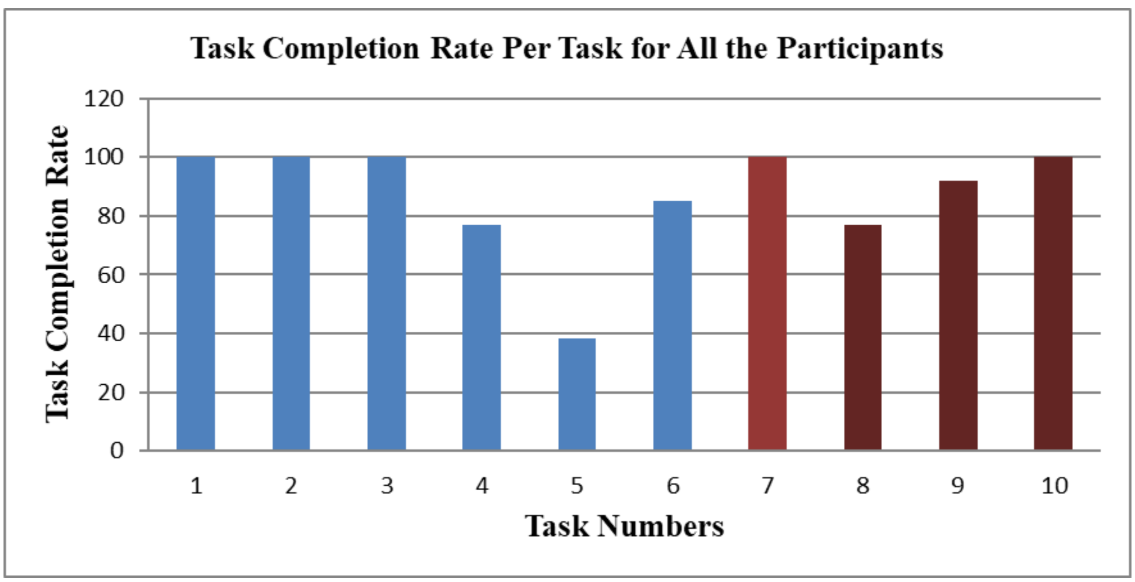


Figure (2). Task Completion Rate Per Task for All the Participants. Blue indicates tasks involving the Area Health Profile (Tasks 1–6); red indicates tasks involving the Double Map (Tasks 7–10)

As Figure (2) shows, four tasks (38.46%) reached the 100% completion success rate, One task (7.69%) got 90% of the completion success rate, One task (7.69%) got 85% or higher completion rate and three tasks (23.08%) got less than 80% of completion success rate. The tasks #1, 2, 3, 7, and 10 got TCR of 100%. As the multi-task scenario shows (see the appendix), the tasks #1, 2, & 10 had very simple context and functionality. The mentioned tasks do not inquire to catch or interpret difficult epidemiology or statistical outcomes. Task # 3 was ranked as a one of the complicated tasks, but it was successfully accomplished by all the subjects.

Task #6 was ranked as one of the complicated tasks, but was accomplished effectively in both study rounds. Task #6 was effectively accomplished by the study subjects in the two rounds of the study, with completion rate of 100% in the first round and just 85% in the second round.

The tasks #4, 5 & 8, had the lowermost TCRs. These tasks were not effectively conducted because the tasks' TCRs were less than 78%. These tasks were ranked as difficult tasks and require certain abilities and understanding to be completed successfully. Some of these tasks are multi-stepping tasks and needs special knowledge to be figured out. From the first round of the study, the same tasks in addition to the task # 3 got the lowest completion rates [17].

Mapping Reports' efficiency and productivity

The investigators used the audio-video records to measure the time per task which was measured from starting the examined task to the time of beginning the next task. The median of task #8 was the highest followed by # 4, 5, 6, and 3. This was relatively connected to the task's difficulty. Table (1) shows the times which our subjects spent on the study tasks individually.

Table (1). Time on the Study Tasks

Task #	Task Time Range	Task Time Mean	Task Time Median
1	6-17 seconds	9.77 seconds	7 seconds
2	6-300 seconds	44.92 seconds	16 seconds
3	6-114 seconds	28.69 seconds	19 seconds
4	9-166 seconds	83.38 seconds	113 seconds
5	12-245 seconds	85 seconds	51 seconds
6	15-196 seconds	82.62 seconds	78 seconds
7	6-43 seconds	11.62 seconds	8 seconds
8	50-429 seconds	182.77 seconds	179 seconds
9	9-26 seconds	13.75 seconds	12 seconds
10	3-7 seconds	5.54 seconds	6 seconds

Time Based Efficiency (TBE) and Overall Relative Efficiency (ORE)

The time based efficiency rates were ranged between 0 goals/second for task #8 and 0.12 goals/second for task #1. Figures (3) and (4) show the time-based efficiency and the overall relative efficiency for every individual performed tasks.

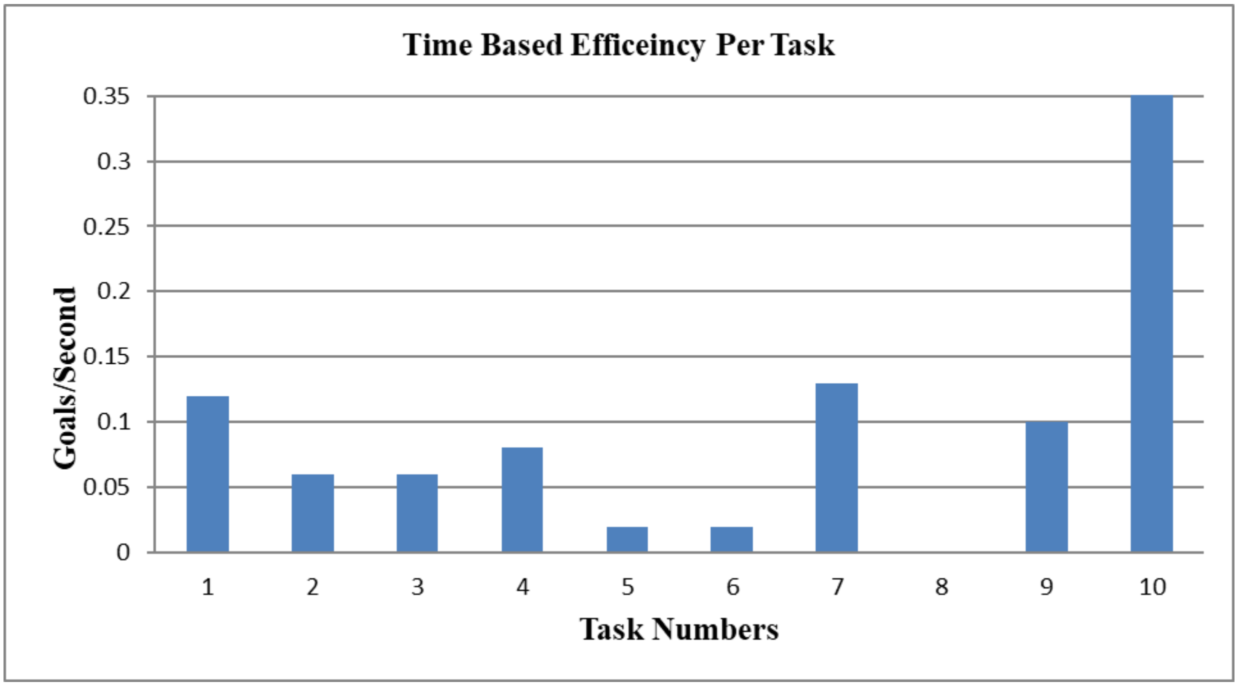


Figure (3). Time-Based Efficiency (TBE) per Task

From Figure (3), the TBE per task was different for the all the tasks. Task #10 gained the maximum TBE (.35 Goals/Second) and this result was associated with ease of the task (close the map), followed by the tasks #7, 9 and 1 and all of them ranked as simple and straight forward tasks (proceed to the "double map" link on the desktop, open the "area profile" map link in the desktop, and check the sources of our mapping report data). Tasks #5, 6, and 8 got the lowest TBE levels, 0 goals/second for the task #8 and .02 goals/second for the tasks # 5 & 6. Tasks #5,6,&8 were ranked as the most complicated tasks.

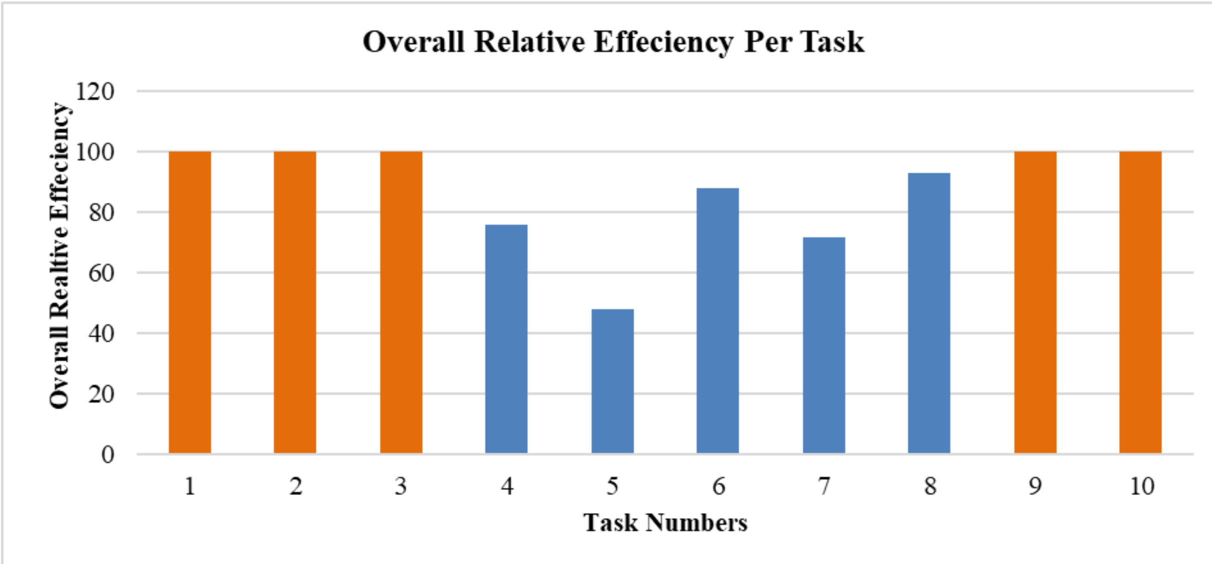


Figure (4). Overall Relative Efficiency per Task. Orange indicates tasks with 100% ORE per Task, blue indicates task with less than 100% ORE per Task

Figure (4) shows that the highest ORE rates were for the tasks #1, 2, 3, 9& 10 and they were categorized as easy tasks except the task #3, which is ranked as a complex task. Task #8 got an OTE of 93% in comparison to almost OTE of 90% in the first round [17]. Task #6, a ranked complicated task, followed with an OTE of 88% ORE rate, in comparison to the OTE of about 95% in the first round [17]. Tasks # 4, 5, & 7 had the lowest ORE per task.

3. Users Satisfaction

The scale was given to every study subject to be completed at the end of the study experiment. This tool was constructed to assess the expectations and the insights of the users regarding the tested systems [20]. Figure (5) presented the study subjects' SUS scores.

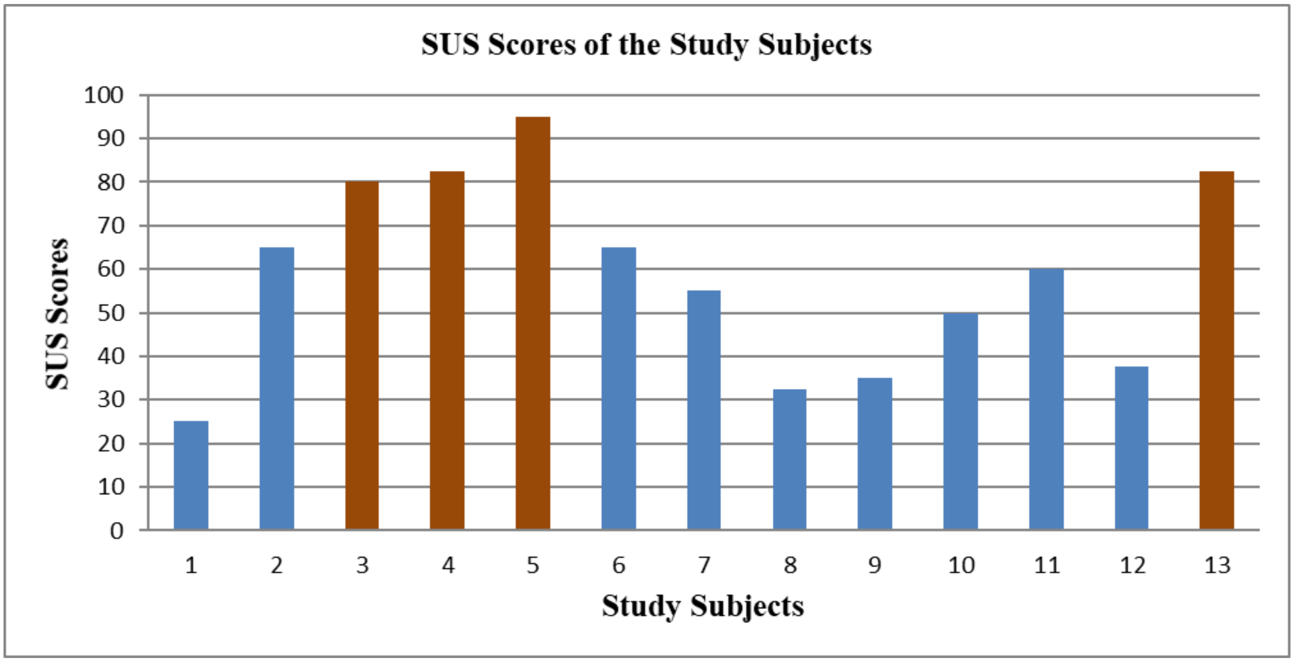


Figure (5). System Usability Scale (SUS) Scores of the Study's Participants. Brown color indicates SUS score of > 68 points, and blue color indicates SUS score of < 68 points

As Figure (5) illustrates, the SUS score range between 0 and 100. The SUS scores for all the current study's subjects stretched from 25 to 82.5 with mean of 58.85 points and median of 65 points. Four of our 13 (30.77%) participants got a SUS score of more than 68 points, and the remaining nine subjects (69.23%) did not reach the accepted satisfaction point, as in the first round of the study [17].

4. Factors affect the participants' performance

Education Level and Work Type Factors

The researchers conducted Wilcoxon-Mann-Whitney test to explore the statistical association among the education level of the participants from one part and the task completion rate and the SUS scale for the study subjects. There was no significant statistical difference between the performance of the graduate school degrees holder participants and the undergraduate degree or less holder participants (P = .72). The results show that there was no significant statistical relationship between the participants' education level and the satisfaction score of the SUS (P = .21).

The same inferential test was also used to assess the statistical relationship between the participants' performance on the test in terms of the task completion rate and the work type of the same participants as cancer researchers and epidemiologists from one part, and participants who do not do research and do not have previous experience in cancer epidemiology or surveillance on the other part. The relationship between the studied two factors was statistically insignificant (P = .63).

All the above results were comparable to the first round's findings [17].

Experience in Healthcare Field and Experience with Mapping Reports and GIS Tools factors

A simple regression test was used to search if there is a significant statistical relationship between the performances of the study participants in the current usability trial and between both: the experience in healthcare field and the previous experience with mapping reports and other GIS tools. The statistical relation between the performances of the participants in the trial and between the experience in healthcare field was insignificant (P = .51). In the current round, there was no significant statistical relationship between the subjects' performance on the study test and the experience in GIS tools (P = .17), while in the previous round the relationship was significant [17].

After we conducted a simple linear regression test, there was no significant statistical relationship between the SUS levels and both the experiment completion rate and the previous experience with GIS tools for the study participants. The P values for these results were (P = .67) and (P = .61) respectively.

Table 2. Demographic and Previous Expertise Factors of the Study Participants Versus the trial's TCR and the Participants' SUS Scores

The Studied Factors	*P*
Education Level vs TCR	.72
Education Level vs SUS Score	.21
Work Type vs TCR	.63
Previous Experience in Healthcare Field vs TCR	.51
Previous Experience in GIS Use vs TCR	.17
Previous Experience in GIS Use vs SUS Score	.61

5. Correlation between the Studied Usability Elements (Effectiveness, Efficiency, and Satisfaction)

We measured the correlation between the major three usability elements, effectiveness, efficiency, and the satisfaction of the users. The current study revealed the following findings:

There was a very weak correlation between the task completion rates and the SUS scores for the participants (r = .31, P = .31), while the correlation was strongly high between these factors in the first round of the study (r = .7, P = .08).

The researchers studied the relation between the task completion rates by task from one side and both the TBE and the ORE factors. The current study discovered that there were strong correlation between the effectiveness and the efficiency usability elements of the studied maps and the correlation was respectively as following (r = .39, *P* = .18) and (r = 81, *P* = .0007). The correlation is weaker than the correlation between these factors in the first round of the study (r = .50, *P* = .25), and (r = .92, *P* = .003).

The current research also discovered a very strong correlation between the time spent by the participants on the given tasks and the study subjects' satisfaction (r = .92, *P* = .000009). This correlation is stronger than the correlation between the same factors in the first round of the study (r = .70, *P* = 0.07) [17].

Table 3. Correlation between the Studied Usability Elements (Effectiveness, efficiency, and Satisfaction)

The Studied Factors	Correlation Coefficient	P
TCR vs SUS Score	.31	.31
TCR vs TBE	.39	.18
TCR vs ORE	.81	.0007
Efficiency Per Participant* vs SUS Score	.92	.000009

*: The total time in seconds of the whole trial per participant

Discussion

Main findings

The current round was conducted to assess the usability of the published mapping report of MCR-ARC using cancer professional participants. This multi-approach usability testing methods might aid map creators to design friendly-used mapping reports and help them to access the maps' prospective users' anticipations. The study findings support using usability testing studies before and after releasing the MCR-ARC maps to the potential users, and support extensive examination of the mapping reports to improve their usability.

Participant Demographics

The researchers conducted this study as a second round of the previous pilot usability study on the same tested mapping reports which are published by MCR-ARC. Previous study includes seven convenient academic health professionals [17]. The investigators refined the tested maps and the usability study's multi-tasks scenario considering all the recommendations and preferences from the first round subjects. The investigators assumed that the first round's results might be tightly connected to the favorites and the insights of the academic health professionals who did not handling cancer data registration and/or analysis, did not directly make and advocate for cancer policy, and did not use cancer mapping reports in their daily practical life [17]. That is why the investigators tried to apply the experiment on the cancer officials who were attending NAACCR-

2016 conference aiming to explore the usability of the refined tested maps using a convenience sample of cancer officials.

Effectiveness and Efficiency

Effectiveness per Participant

In the current round, we are aiming to attain 100% of completion success rate per participant, but the usability specialists allocated that for any usability study it is ok to consider 78% as the average TCR per participant [24].

According to the results in Figure (1), the trial was completed effectively by twelve out of our 13 participants (92.30%). The trial was conducted effectively despite the diversity in the education, public health field experience, or GIS experience of the study subjects. A PhD holder subject with a heavy cancer and public field experience could not accomplish the minimum border of the accepted TCR, while the other lower educated and less experienced subjects could handle the test effectively.

Effectiveness per Task

As the results section shows, the ranked easy tasks were accomplished effectively than the tasks were ranked as complicated. These findings supported the previous study's round and a previous scientific study's results, which revealed that tasks accomplishments are influenced by the task context's complexity [17,29].

Surprisingly, task #6 was graded as one of the complex tasks but was conducted successfully in both study rounds [17]. In this study round, it could be explained that because the task is very linked to the prior tasks and it was easy to handle after the subject solve the former tasks. The variance in TCRs per task between the two study's rounds might be interpreted as following: The second round's participants' educational backgrounds were very diverse and not all of them had solid epidemiological and/or statistical background that the first round participants hold.

Efficiency

As in the first round of the study, there was difference in the time per task even with the TCR of 100%. The study subjects in the two study rounds spent various time to accomplish the given tasks. As we revealed from the first round of the study, these findings were relatively related to the complexity of the tasks, where the complex tasks toke longer time than the easy ranked ones [17].

Also as we revealed from the first round, this round of study discovered that even the simply ranked tasks, some of the subjects consumed longer time to accomplish the tasks effectively than the other subjects [17]. These results are unpredictable because, as we discussed previously, the study investigators adjusted the tested maps according to the first round subjects' comments. We were looking for making the tasks performed by all the users within comparable times. Based on the TBE results from the previous study round and the current one, the investigators expected that in addition to the intellect and knowledge which are essential to achieve these tasks the usability problems also could even make the given tasks more difficult than the study researchers assumed before they started the study [17]. For both study rounds, we revealed that OTEs for most tasks were fairly related to the difficulty of these tasks' context, and this supports the previous usability literature findings [17,22]. But surprisingly, this is inapplicable for the previously ranked complicated tasks, #3, 6 & 8. After close scrutiny of the recordings, we discovered that the repeatability and re-doing of the preceding tasks profoundly influenced the tasks #3, 6 and 8 completion by the study subjects in both study rounds.

Participants satisfaction

Sixty eight points or more has been considered satisfactory according to usability literature and advanced marks through 100 points exemplifies the optimal to greatest SUS score. In comparison to the first round of the study, we found that both rounds have an averages and the medians measures of less than 68 points and were closed to each other [17].

Our explanation to the comparability of the satisfaction results between both study rounds is that in the second round, we tested cancer professionals of more varied races and with mixed graduate and undergraduate levels and most of these participants had null previous experience in using GIS. These participants developed a comparable satisfaction results as the first round study's subjects, who were holding graduate degrees and had rich experience in statistical and epidemiological knowledge, as well as previous experience in using GIS tools [17]. We assumed that, when we updated our maps according to the first round's results, we simplified our tested maps to fit the needs of our potential users of different biographic and experience levels.

Usability scholars revealed that the users' insights and anticipations are critical to building faultless technology. Considering all the users' commentaries are essential to make the technology more satisfactory and useful [30]. We considered the participants' texts which were taken at the end of every trial, in addition to the audio-visual records of all the participants to explore more usability issues of the tested maps. The revealed usability issues helped us to explain and find out why some of the assigned study tasks were hard to be handled by the highly educated and knowledgeable subjects. The study researchers considered all of the discovered usability problems, and accordingly, will refine our published mapping reports and publicize them on the registry's website.

Factors affect the participants' performance

There weren't significant statistical relationships between all the studied factors. While there was a significant statistical relationship between TCR rates and the previous experience in using GIS tools in the first round of the study, in this round, there is no significant statistical relationship between these two factors. The current study's finding might be explained as following: By considering the round one's results and by updating our maps, we made the tested maps simpler to be used by the subjects who had null experience in using GIS technology. Both study rounds revealed that there is no dependency of the SUS score on the both the TCR and the previous experience with GIS tools for the study participants as the investigators assumed [17].

The Correlation between the Studied Usability Elements (Effectiveness, Efficiency, and Satisfaction)

The results revealed positive correlation between the three usability elements which range from very weak to very strong positive correlations. The strong correlation between the SUS and the efficiency is very reasonable, and this supported the findings of the previous round [17]. This could be explained that by updating the maps, we made the maps more usable by the users and the participants conducted the trial more efficiently and they were satisfied about the entire experience.

Strength of the study

This study is the second round of the usability research conducted on the published MCR-ARC InstantAtlas mapping reports to evaluate the quantitative and qualitative measurements using a larger sample than we used for the first round [17]. The investigators used a sample of 13 participants who are professionals in cancer epidemiology, cancer surveillance and/or cancer research. We are assuming that the second round professional participants will be the potential users of the interactive mapping reports than the academia people who had been selected for the first round trial. Larger sample of cancer professionals and testing the modified reports based on the first round's findings, make this round more reasonable and the results tend to be more applicable [17]. These results might be generalized to assess the usability and the functionality of all the MCR-ARC's mapping reports.

Conclusion

Current round of the study measured the three main components of the refined tested mapping reports: effectiveness, efficiency, and user satisfaction. The study results supported the first round's findings that the three usability elements are correlated positively to each other. As we revealed from the previous round, the examined reports' effectiveness results were superior in comparison to the efficiency and satisfaction results. As we pointed out from the first round, we revealed that the effectiveness and efficiency metrics were strongly associated with the trial tasks context's complexity. As we concluded previously in the previous study, the graded simple tasks were achieved effectively and efficiently easier than the graded complicated tasks. The SUS scores of the current study round were comparable to the previous round study's SUS scores with a poor average and median while the SUS score ranges of the both studies were ranged from very poor to excellent scores [17].

The current study, as the opposite of the previous round, showed that there was no significant statistical relationship between the subjects' performance on the study test and having previous experience in using the GIS tools [17]. Updating the tested maps and tasks, made the reports simpler to be used even by users who do not have previous GIS experience.

In a nutshell, the mapping reports must be extensively refined and modified to correct all the revealed usability concerns and to meet the perceptions and requirements of the maps' potential users. Also, the two round study methodology could be applied on other MCR-ARC atlases, and might be served to improve the usability of these maps. Including the GIS tools' users should be considered at the initial phases of scheming and creating the GIS reports.

Limitations of the Study

There are several limitations for the study. The sampling technique was the convenience sampling method. The methodology might not determine all the detailed performance and behavioral usability metrics of the participants. The audio-video records were analyzed manually by the primary investigator. We might also think of introducing more sophisticated methodology using advanced usability software, for example, advanced usability software and/or Eye Tracking software to record and analyze our usability data.

Future Directions and Recommendations

We are recommending conducting usability testing pilots on all current and the future designed mapping reports by the MCR-ARC. Also, we are targeting that usability assessment should begin from the planning procedure, map release, and after dissemination the mapping reports by using a representative sample of these maps' probable users to precisely assess the usability of these maps. Additionally, we are hoping that we could use more advanced usability tools in future to evaluate our released maps.

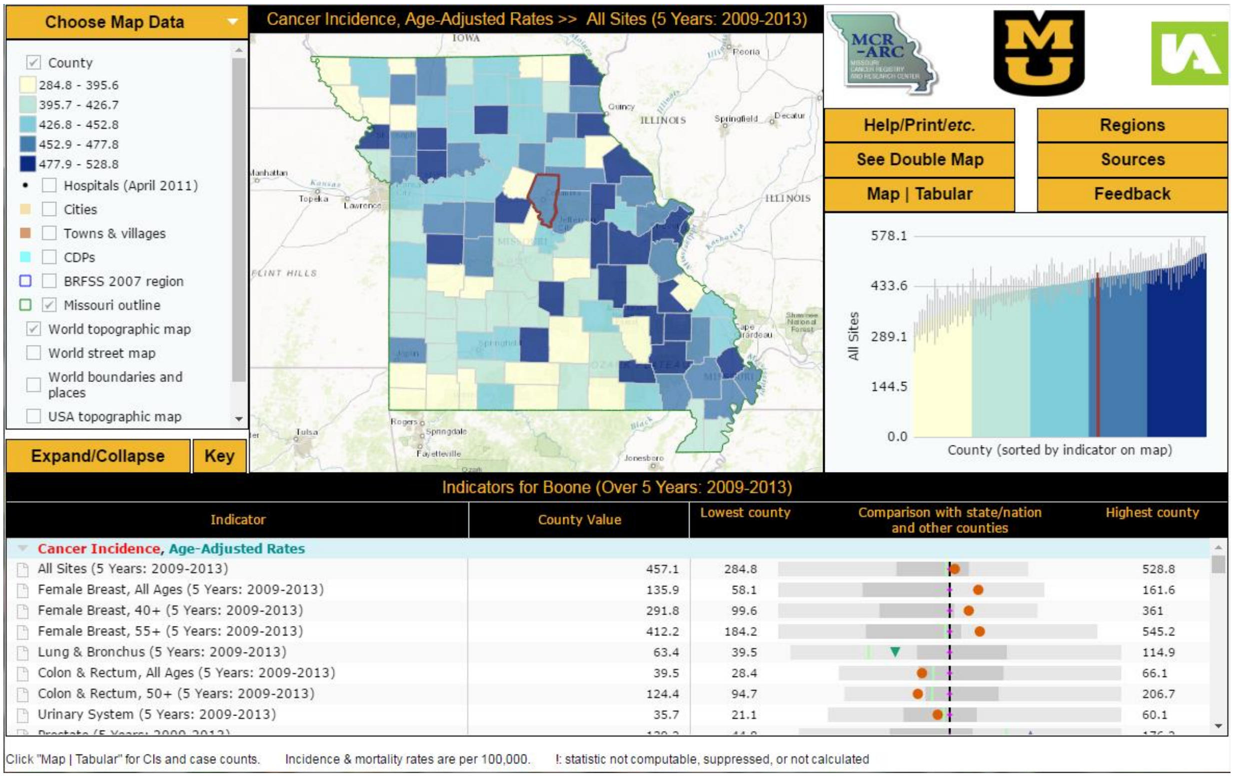
Acknowledgements

Core activities of the Missouri Cancer Registry is supported in part by a cooperative agreement between the Centers for Disease Control and Prevention (CDC) and the Missouri Department of Health and Senior Services (DHSS) (5NU58DP003924-04/05) and a Surveillance Contract between DHSS and the University of Missouri. The publication's content is exclusively the responsibility of the authors and does not necessarily reflect the views of the funders.

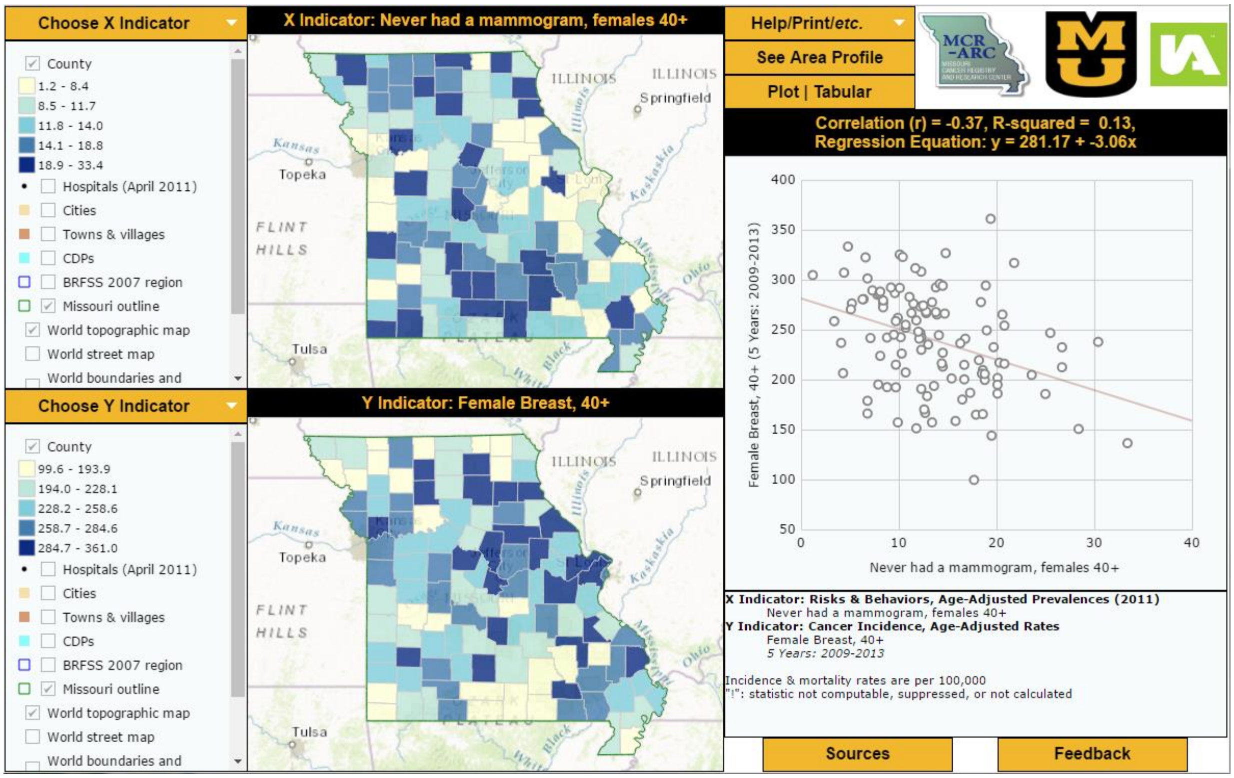
Conflicts of Interest

None declared

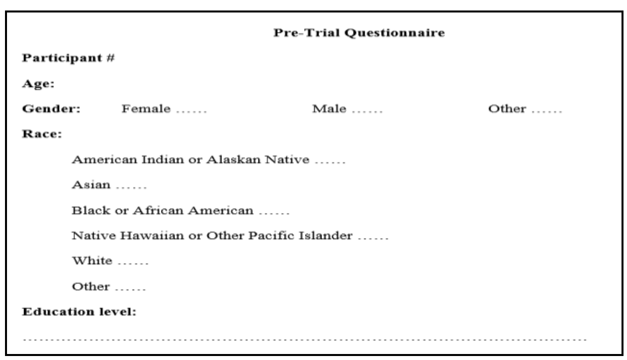
Multimedia Appendix 1



Multimedia Appendix 2



Multimedia Appendix 3



Pre-Trial Questionnaire

Participant #

Age:

Gender: Female ...... Male ...... Other ......

Race:

American Indian or Alaskan Native ......

Asian ......

Black or African American ......

Native Hawaiian or Other Pacific Islander ......

White ......

Other ......

Education level:

....

Multimedia Appendix 4

Multi-Tasks Usability Testing

Task (1):

Open the "**Area Health Profile**" link on the desktop

Task (2):

Open "Choose Map Data"

Task (3):

Choose the age-adjusted incidence rates for Female Breast 55+ for 2008-2012

Task (4):

Write down the incidence rate and the 95% confidence interval (CI) for Boone County

Rate: .... **Confidence Interval (CI)**..................................

Task (5):

Was the rate statistically significantly higher, lower, or no different from the statewide rate?

....

References

1. Haklay M, Tobón C. 2003. Usability evaluation and PPGIS: towards a user-centred design approach. *Int J Geogr Inf Sci*. 17(6), 577-92. doi:https://doi.org/10.1080/1365881031000114107.

2. Brooks Jr FP. The Mythical Man-Month: Essays on Software Engineering, Anniversary Edition, 2/E. Pearson Education India; 1995 Sep 1. ISBN-13: 978-0201835953 ISBN-10: 0201835959

3. Young RR. Effective requirements practices. Addison-Wesley Longman Publishing Co., Inc.; 2001 Mar 1. ISBN:0-201-70912-0

4. Bell BS, Hoskins RE, Pickle LW, Wartenberg D. 2006. Current practices in spatial analysis of cancer data: mapping health statistics to inform policymakers and the public. *Int J Health Geogr*. 5(1), 49. doi:https://doi.org/10.1186/1476-072X-5-49. PubMed

5. Lipkus IM, Samsa G, Rimer BK. 2001. General performance on a numeracy scale among highly educated samples. *Med Decis Making*. 21(1), 37-44. doi:https://doi.org/10.1177/0272989X0102100105. PubMed

6. Sheridan SL, Pignone M. 2002. Numeracy and the medical student's ability to interpret data. Effective clinical practice. *ECP*. 5(1), 35-40. PubMed

7. Woloshin S, Schwartz LM, Moncur M, Gabriel S, Tosteson AN. 2001. Assessing values for health: numeracy matters. *Med Decis Making*. 21(5), 382-90. doi:https://doi.org/10.1177/0272989X0102100505. PubMed

8. Higgs G, Smith DP, Gould MI. 2005. Findings from a survey on GIS use in the UK National Health Service: organisational challenges and opportunities. *Health Policy*. 72(1), 105-17. doi:https://doi.org/10.1016/j.healthpol.2004.06.011. PubMed

9. MacEachren AM. How maps work: representation, visualization, and design. Guilford Press; 2004.Frye C. Making maps that communicate. ArcUser. 2001;4:38-43. ISBN 0-89862-0 (hc.) ISBN 1-5730-040-X (pbk.)

10. Frye C. Making maps that communicate. ArcUser. 2001;4:38-43. http://www.esri.com/news/arcuser/1001/files/bettermaps.pdf. Archived at: http://www.webcitation.org/6mjN4934C

11. Richards TB, Croner CM, Rushton G, Brown CK, Fowler L. 1999. Information technology: Geographic information systems and public health: Mapping the future. *Public Health Rep*. 114(4), 359. PubMed https://doi.org/10.1093/phr/114.4.359

12. Pickle LW. Atlas of United States mortality. National Center for Health Statistics, Centers for Disease Control and Prevention, US Dept. of Health and Human Services; 1996. https://www.cdc.gov/nchs/data/misc/atlasmet.pdf. Archived at http://www.webcitation.org/6mjRxC5qv

13. Paolino L, Sebillo M, Cringoli G. 2005. Geographical information systems and on-line GIServices for health data sharing and management. *Parassitologia*. 47(1), 171. PubMed

14. Bhowmick T, Robinson AC, Gruver A, MacEachren AM, Lengerich EJ. 2008. Distributed usability evaluation of the Pennsylvania Cancer Atlas. *Int J Health Geogr*. 7(1), 36. doi:https://doi.org/10.1186/1476-072X-7-36. PubMed

15. Bell BS, Hoskins RE, Pickle LW, Wartenberg D. 2006. Current practices in spatial analysis of cancer data: mapping health statistics to inform policymakers and the public. *Int J Health Geogr*. 5(1), 49. doi:https://doi.org/10.1186/1476-072X-5-49. PubMed

16. Richards TB, Berkowitz Z, Thomas CC, Foster SL, Gardner A, et al. 2010. Choropleth map design for cancer incidence, part 1. *Prev Chronic Dis*. 7(1):A23.

17. Ben Ramadan AA, Jackson-Thompson J, Schmaltz CL. 2017. Usability Assessment of the Missouri Cancer Registry's Published Interactive Mapping Reports: Round One. *JMIR Human Factors*. 4(3). doi:https://doi.org/10.2196/humanfactors.7899. PubMed

18. Missouri Cancer Registry. MCR-ARC: Cancer Demographics and Behavioral Risks, Area health profile report. 2016. https://instantatlas.umh.edu/IAS/DynamicReports/areaHealthProfile2015/atlas.html?select=019. Archieved at: http://www.webcitation.org/6mkkdZuFX

19. Missouri Cancer Registry. MCR-ARC: Cancer Demographics and Behavioral Risks, Double map report. 2016. https://instantatlas.umh.edu/IAS/DynamicReports/doubleMap2015/atlas.html?indicator2=i2 Archived at: http://www.webcitation.org/6msJd7tzw

20. Brooke J. SUS-A quick and dirty usability scale. Usability evaluation in industry. 1996 Sep;189(194):4-7. ISBN 0748403140 (cased) ISBN 07484404600 (Paperback)

21. Demir F, Karakaya M, Tosun H. 2012. Research methods in usability and interaction design : Evaluations and case studies, 2nd ed. LAP LAMBERT Academic Publishing, Germany. ISBN:3847343351 9783847343356

22. Faulkner L. 2003. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behav Res Methods Instrum Comput*. 35(3), 379-83. PubMed https://doi.org/10.3758/BF03195514

23. Tullis, T., Albert B. Issues-Based metrics, Measuring the user experience: collecting, analyzing, and presenting usability metrics. Newnes; 2013 May 23: 63-119. ISBN 978-0-12-415781-1

24. ISO/TS satge 20282-2(en): Usability of consumer products and products for public use-Part 2: Summative test method, 2013, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/obp/ui/#iso:std:iso:ts:20282:-2:ed-2:v1:en Archived at: http://www.webcitation.org/6mnzjUyzL

25. Mifsud J. Usability Metrics- A guide to quantify the usability of any system. Usability Geek. 2015June22. http://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/. Archived at: http://www.webcitation.org/6pfgKs9MK

26. Clarke MA, Belden JL, Kim MS. 2014. Determining differences in user performance between expert and novice primary care doctors when using an electronic health record (EHR). *J Eval Clin Pract*. 20(6), 1153-61. PubMed https://doi.org/10.1111/jep.12277

27. Marx A, Backes C, Meese E, Lenhof HP, Keller A. 2016. EDISON-WMW: Exact Dynamic Programing Solution of the Wilcoxon–Mann–Whitney Test. *Genomics Proteomics Bioinformatics*. 14(1), 55-61. doi:https://doi.org/10.1016/j.gpb.2015.11.004. PubMed

28. Schoenfeld D. Statistical considerations for a study of the effect of one variable on another, June 8 2017. http://hedwig.mgh.harvard.edu/sample_size/js/js_associative_quant.html Archived at: http://www.webcitation.org/6r4aHpevO

29. Bhowmick T, Gruver A, Robinson AC, MacEachren AM, Lengerich E, et al. Using e-delphi to evaluate the Pennsylvania cancer atlas. InAutoCarto 2006: Proceedings of the AutoCarto 2006 Conference, CAGIS, Vancouver, WA, ESRI, Redlands, CA 2006 Jun 26. https://www.researchgate.net/profile/Anthony_Robinson6/publication/250050952_Usin_eDelphi_to_Evaluate_the_Pennsylvania_Cancer_Atlas/links/54d7948e0cf2970e4e745de.pdf. Archived at: http://www.webcitation.org/6r30uTME8

30. El-Abed M, Giot R, Hemery B, Rosenberger C. A study of users' acceptance and satisfaction of biometric systems. InSecurity Technology (ICCST), 2010 IEEE International Carnahan Conference on 2010 Oct 5 (pp. 170-178). DOI: 10.1109/CCST.2010.5678678