EHR Usability Test Report of Soteria vs 6.1

Customized Common Industry Format Template for Electronic Health Record Usability Testing Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Report Prepared By: Chart Lux Consulting

Kyle Meadors

President

(615) 804 9600

kyle@chartlux.com

7333 Riverfront Dr.

Nashville, TN 37221

Date of Usability Test: October 15-November 15, 2023

Date of Report: December 03, 2023

Table of Contents

EXECUTIVE SUMMARY

A usability test of Soteria vs 6.1 was conducted virtually during October 15-November 15, 2023 by Chart Lux Consulting. The purpose of this test was to test and validate the usability of the current user interface and provide evidence of usability in the EHR Under Test (EHRUT). During the usability test, ten (10) healthcare providers and individuals matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

This study collected performance data on four (4) tasks typically conducted on an EHR for documenting demographics:

- Record, Change, and Access Demographics
- Clinical Decision Support
- CPOE Diagnostic Labs
- Record and Change Implantable Devices

During the 60 minute one-on-one usability test, each participant was greeted by the administrator, and they were instructed that they could withdraw at any time (included in Appendix B). Healthcare, and participants had no direct connection to the development of the EHRUT. The participants had not used any version of this EHRUT before the test event. Participants received a brief training and orientation of new features prior to testing.

The administrator introduced the test and instructed participants to complete the task using the EHRUT. During the testing, the proctor timed the test and recorded user performance data on paper and electronically. The administrator did not give the participant assistance in how to complete the task. Participant screens, head shots and audio were recorded for subsequent analysis.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Following the conclusion of the testing, participants were asked to complete a post-test questionnaire and were compensated for their time. Various recommended metrics, in accordance with the examples set forth in the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, were used to evaluate the usability of the EHRUT. Following is a summary of the performance and rating data collected on the EHRUT.

Measure	N	Task Success	Path Deviations	Task Time (Seconds)		Errors	Task Ratings (5=Easy)
Task	#	Mean (SD)	Deviations (Observed / Optimal)	Mean (SD)	Deviations (Observed / Optimal)	Mean (SD)	Mean (SD)
Record, Change, and Access Demographics	10	100% (0%)	12 / 10	114 (37)	37 / 60	10% (30%)	4.9 (0.3)
Clinical Decision Support	10	95% (22%)	23 / 15	243 (83)	83 / 75	3% (10%)	4.2 (0.6)
CPOE Diagnostic Labs	10	85% (36%)	9 / 5	54 (22.4)	22.4 / 30	0% (0%)	4.9 (0.3)
Record and Change Implantable Devices	10	85% (36%)	11 / 5	90 (67)	67 / 4	0% (0%)	4.2 (1.2)

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 91. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

Major Findings

Participants gave the system high marks and noted it to be very usable and praised its simplicity, intuitiveness, and ease of use. They appreciated the ability to see all key information on the same screen of view to perform the tasks at hand.

Areas for Improvement

While results were good and high marks given, some comments were made that extra button clicks were needed to save the chart or the updated field when they would prefer the result be automatically saved after leaving the respective field. Also, some comments were made no wishing the font size was slightly larger. We will take this feedback into account for future design considerations.

Usability Report

Introduction

The EHRUT tested for this study was Soteria vs 6.1, designed to support healthcare providers treating patients in ambulatory settings. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency and user satisfaction, such as time to complete the tasks and deviations from optimal pathways, were captured during the usability testing.

Method

Design Standard

Soteria employed NISTIR 7741 usability standard in our product design. It is a user-centered design (UCD) created for improving the usability of electronic health records (https://www.nist.gov/publications/nistir-7741-nist-guide-processes-approach-improving-usability-electronic-health-records).

Participants

A total of ten (10) participants were tested on the EHRUT. Participants in the test primarily act in the role of physicians and physician assistants. Participants were recruited by Infocare Healthcare Systems. Healthcare, and participants had no direct connection to the development of the EHRUT. The participants had not used any version of this EHRUT before the test event. Participants received a brief training and orientation of new features prior to testing.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

ID	Gender	Age	Education	Role	Prof Experience (months)	Comp Experience (months)	Product Experience (months)	Assistive Technology Needs
01	F	30-39	High school graduate	partner	180	180	0	None
02	F	30-39	Bachelor's degree	nurse	132	132	0	None

ID	Gender	Age	Education	Occupation				
03	F	30-39	Doctorate degree	physician	36	36	0	None
04	F	30-39	Master's degree	nurse	144	144	0	None
05	F	40-49	Some college credit, no degree	office manager	240	240	0	None
06	F	50-59	Bachelor's degree	nurse	384	300	0	None
07	F	40-49	Bachelor's degree	trainer	120	120	0	None
08	M	50-59	Doctorate degree	nurse practitioner	300	240	0	None
09	F	20-29	Bachelor's degree	nurse	60	60	0	None
10	M	30-39	Doctorate degree	clinical director	120	144	0	None

All ten participants (matching the demographics in the section on Participants) were recruited and participated in the usability test. Participants were scheduled for 60 minute sessions with the test screener.

Study Design

Overall, the objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors

- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the Usability Metrics section.

Tasks

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR according to its respective ONC certified criteria. Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Tasks used in the study are listed below and with their relative risk associated with user errors noted.

1. Record, Change, and Access Demographics (Low Risk)
2. Clinical Decision Support (Medium Risk)
3. CPOE Diagnostic Labs (High Risk)
4. Record and Change Implantable Device (Low Risk)

Procedures

Test participants were scheduled for 60 minute sessions and arrived as individual participants. Each participant was assigned a number to identify results while detaching the identity of the individual from the response and observations. Demographic data was collected from each participant matched with a name on the participant schedule.

A test administrator moderated each test including administering instructions and tasks. The administrator also monitored path deviations and task success, obtained post-task rating data, and took notes on participant comments. The test administered monitored task times and took notes on number and types of errors, using the recorded video session to confirm details.

Participants were instructed to perform the tasks as quickly as possible, making as few errors and deviations as possible, and without assistance.

Each participant was provided with a clinical scenario providing the background context for the task workflows. Each participant had the scenario read to them and then received training on how to perform the task. Task timing began once the participant began the task. The task time was stopped once the participant successfully completed the task. Scoring is discussed below.

Following the session, the administrator gave the participant the post-test questionnaire on usability (see Appendix D), provided instructions on how compensation for their time would occur, and thanked each individual for their participation.

Test proctor compiled the demographic information, task success rate, time on task, errors, deviations, comments, and post-test questionnaire for analysis and scoring.

Test Location

Testing was done using Teams or Zoom remote session. Only one participant was logged in at any given time with the administrator to ensure privacy.

Test Environment

The EHRUT would typically be used in an ambulatory setting, and the testing environment was set up to mimic this workflow. The test application was running on a private server using a test database on an Internet connection. The participants used a mouse and keyboard when interacting with the EHR.

The application was set up by Infocare Healthcare Systems engineering to mimic a live environment. Technically, the system performance (i.e. response time) was representative of what actual users would experience in a field implementation. Additionally, participants were not allowed to change any of the default system settings.

Test Forms and Tools

During the usability test, various documents and instruments were used, including:

1. Demographics Questionnaire
2. Participant Briefing/Debriefing document
3. Usability Task Tracking document
4. Post-Test Questionnaire (System Usability Scale)

Examples of these documents can be found in Appendices A-D respectively.

The participant's interaction with the EHRUT was captured and recorded digitally with web conferencing software running on the test machine. The test administrator participated in each session live, with access to the recorded session afterwards.

Participant Instructions

The administrator reads the following instructions noted in Appendix B. Participants were given four (4) tasks to complete. Tasks are listed in the Usability Task Tracking document in Appendix C.

Usability Metrics

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an

acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing.

The goals of the test were to assess:

1. Effectiveness of Soteria by measuring participant success rates and errors
2. Efficiency of Soteria by measuring the average task time and path deviations
3. Satisfaction with Soteria by measuring ease of use ratings

Data Scoring

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rational and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor (e.g. 1.25) that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was 60 seconds then allotted task time performance was 80 seconds (60 x 1.25). This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as an "Failures." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency:	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control.

Task Deviations	This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the Soteria overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix D.

Results

Data Analysis and Scoring

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. There were no participants who failed to follow session and task instructions and as a result all participants had their data included in the final analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that, if corrected, yield material, positive impact on user performance.

Measure	N	Task Success	Path Deviations	Task Time (Seconds)		Errors	Task Ratings (5=Easy)
Task	#	Mean (SD)	Deviations (Observed / Optimal)	Mean (SD)	Deviations (Observed / Optimal)	Mean (SD)	Mean (SD)
Record, Change, and Access Demographics	10	100% (0%)	12 / 10	114 (37)	37 / 60	10% (30%)	4.9 (0.3)
Clinical Decision Support	10	95% (22%)	23 / 15	243 (83)	83 / 75	3% (10%)	4.2 (0.6)
CPOE Diagnostic Labs	10	85% (36%)	9 / 5	54 (22.4)	22.4 / 30	0% (0%)	4.9 (0.3)
Record and Change Implantable Devices	10	85% (36%)	11 / 5	90 (67)	67 / 4	0% (0%)	4.2 (1.2)

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 91. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

Discussions of the Findings

Effectiveness

Based on the success, failure and path deviation data, the system was well designed, and the tasks scored well.

Efficiency

Efficiency was measured as a function of time on task relative to pre-determined benchmark task times and clicks per task relative to benchmark task values. Based on the task completion times, the majority of users completed all tasks close to the optimal time. A few participants needed to refer back multiple times to the test story to enter in the correct information which delayed completion of the task.

Satisfaction

Based on the task ratings, all the participants found the tasks to be intuitive and easy to perform. The SUS score was calculated to be 94, with the lowest score being 80 and the highest score being 100.

Major Findings

Participants gave the system high marks and noted it to be very usable and praised its simplicity, intuitiveness, and ease of use. They appreciated the ability to see all key information on the same screen of view to perform the tasks at hand.

Areas for Improvement

While results were good and high marks given, some comments were made that extra button clicks were needed to save the chart or the updated field when they would prefer the result be automatically saved after leaving the respective field. Also, some comments were made no wishing the font size was slightly larger. We will take this feedback into account for future design considerations.

Appendix A: Demographic Questionnaire

Name	
Gender	
Age	
Education (highest attained)	
Clinical Role	
Professional Experience (in months)	
Experience with Computers in Healthcare (in months)	
Experience with EHR (in months)	

Appendix B: Participant Briefing/Debriefing

Thank you for participating in this study. Our session today will last approximately 30 minutes. During that time, you will look at our EHR and be asked to do various tasks associated with its ONC certification criteria. The goal is for you to attempt to complete the various tasks to the best of your ability, and we will document your findings as part of our effort to certify our product in the ONC health IT certification program.

The product you will be using today is not ready for production, but the functionality you will be encountering in the testing tasks is nearly at its finish state for this upcoming release. While we provide a clinical story for the test tasks at hand, some of the test data we provide may not make sense for your personal day-to-day activities and it should be treated as placeholder data for testing.

I will ask you to complete a few tasks using this system and then answer some questions. We are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. You will be asked to complete these tasks on your own trying to do them as quickly as possible with the fewest possible errors or deviations. Do not do anything more than asked. If you get lost or have difficulty, I cannot answer help you with anything to do with the system itself. Please save your detailed comments until the end of a task or the end of the session as a whole when we can discuss freely. Please be honest with your opinions as this feedback will help improve the product.

We are recording our session today via web conferencing software. All of the information that you provide will be kept confidential and your name will not be associated with your comments at any time.

Do you have any questions or concerns?

Appendix C: Usability Tasks

Task 1: Record, Change, and Access Demographics

1. Take the participant to the starting point for the task. Begin timer.
2. User will perform the actions according to the assigned patient data sheet.
3. Record Success:
    a. Completed according to proper steps.
    b. Completed with difficulty or help. Describe below in comments.
    c. Not completed.
    d. Comments:
4. Task Time Observed (seconds):
5. Task Time Optimal (seconds):
6. Optimal Path:
    a. Start at Scheduler
    b. Select registered Patient
    c. Go to Patient Details
    d. Select Edit.
    e. Update DOB to 01/19/1956
    f. Update Sex to Male
    g. Update Race to Black
    h. Update Ethnicity to Hispanic/Latino
    i. Record Preferred Language as English
    j. Record Gender Identity to Male
    k. Record Sexual Orientation as Choose not to disclose
7. Results of Pathway Choice
    a. Correct
    b. Minor Deviations/Cycle (describe below)
    c. Major Deviations (describe below)
    d. Comments:
8. Record Errors and Verbalizations:
9. Ask participant: "overall, how would you rate this task? Rating: (5) Very Easy (4) Easy (3) Moderate (2) Difficult (1) Very Difficult":

Associated Criteria:

▪ 170.315(a)(5) Demographics

Task 2: Clinical Decision Support

1. Take the participant to the starting point for the task. Begin timer.
2. User will perform the actions according to the assigned patient data sheet.
3. Record Success:
    a. Completed according to proper steps.
    b. Completed with difficulty or help. Describe below in comments.
    c. Not completed.
    d. Comments:
4. Task Time Observed (seconds):
5. Task Time Optimal (seconds):
6. Optimal Path:
    a. Start at Daily Census
    b. Start at Patient Chart
    c. Select a new clinical note.
    d. Choose master template for new note.
    e. View alerts in Chart.
    f. Go to Orders.
    g. Select Imaging.
    h. Record mammogram screening.
    i. Add problem.
    j. Add medication.
    k. Add allergy.
    l. Save note.
    m. Re-open note.
    n. View alerts in chart.
7. Results of Pathway Choice
    a. Correct
    b. Minor Deviations/Cycle (describe below)
    c. Major Deviations (describe below)
    d. Comments:
8. Record Errors and Verbalizations:
9. Ask participant: "overall, how would you rate this task? Rating: (5) Very Easy (4) Easy (3) Moderate (2) Difficult (1) Very Difficult":

Associated Criteria:

▪ 170.315(a)(9) Clinical Decision Support

Task 3: CPOE Diagnostic Labs

1. Take the participant to the starting point for the task. Begin timer.
2. User will perform the actions according to the assigned patient data sheet.
3. Record Success:
    a. Completed according to proper steps.
    b. Completed with difficulty or help. Describe below in comments.
    c. Not completed.
    d. Comments:
4. Task Time Observed (seconds):
5. Task Time Optimal (seconds):
6. Optimal Path:
    a. Start on Patient Chart.
    b. Edit previously created note.
    c. Click on Orders (bottom right).
    d. Select In House.
    e. Select Labs and record data assigned in story.
    f. Add labs.
    g. Save note.
    h. Re-open note.
    i. Confirm lab is saved.
7. Results of Pathway Choice
    a. Correct
    b. Minor Deviations/Cycle (describe below)
    c. Major Deviations (describe below)
    d. Comments:
8. Record Errors and Verbalizations:
9. Ask participant: "overall, how would you rate this task? Rating: (5) Very Easy (4) Easy (3) Moderate (2) Difficult (1) Very Difficult":

Associated Criteria:

▪ 170.315(a)(2) CPOE Labs

Task 4: Record and Change Implantable Devices

1. Take the participant to the starting point for the task. Begin timer.
2. User will perform the actions according to the assigned patient data sheet.
3. Record Success:
    a. Completed according to proper steps.
    b. Completed with difficulty or help. Describe below in comments.
    c. Not completed.
    d. Comments:
4. Task Time Observed (seconds):
5. Task Time Optimal (seconds):
6. Optimal Path:
    a. Start at Patient Chart.
    b. Select Clinical Notes
    c. Edit previously created note.
    d. Select Past Medical, Surgical, Personal/Social and Family History
    e. Add Surgical History
    f. Enter procedure name.
    g. Enter UDI.
    h. Enter Medical Device Name.
    i. Set status as Insitu.
    j. Select Device Details
    k. Select Add
    l. Save the note as draft
    m. Re-open note.
    n. Confirm UDI is saved.
7. Results of Pathway Choice
    a. Correct
    b. Minor Deviations/Cycle (describe below)
    c. Major Deviations (describe below)
    d. Comments:
8. Record Errors and Verbalizations:
9. Ask participant: "overall, how would you rate this task? Rating: (5) Very Easy (4) Easy (3) Moderate (2) Difficult (1) Very Difficult":

Associated Criteria:

▪ 170.315(a)(14) Record and Change Implantable Devices

Appendix D: System Usability Scale

Ratings: Strongly Agree (5)   Agree (4)   Neutral (3)   Disagree (2)   Strongly Disagree (1)

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.