EHR Usability Test Report of the Chirp EHR (Version 1.0)

Report based on NISTIR 7742 Common Industry Format for Usability Test Reports

Dates of Usability Test: November 5 to November 10, 2020 and December 9 to 11
Date of Report: December 14, 2020

Report Prepared By:

The Usability People, LLC
4000 Legato Road, Suite 1100
Fairfax, VA 22033

www.TheUsabilityPeople.com

TABLE OF CONTENTS

Executive Summary

During the time period of November 5 to November 10, 2020 and December 9 to 11, 2020, The Usability People, LLC conducted a summative usability test of the Chirp EHR. The test was conducted in the Fairfax, VA office of The Usability People over individual remote tele-conferencing sessions. The purpose was to test and validate the usability of the current user interface and provide evidence of usability of the Chirp EHR as the EHR Under Test (EHRUT). Thirteen (13) healthcare providers matching the target demographic criteria participated in the usability test using the EHRUT in simulated, but representative tasks.

The study focused on measuring the effectiveness of, efficiency of, and satisfaction (ISO 9241-11) with the Chirp EHR among a sample of participants representing potential users of the system. Performance data was collected on thirteen (13) tasks typically conducted on an EHR. Tasks created were based upon the criteria specified within the test procedure structure for evaluating conformance of Electronic Health Record (EHR) technology to the certification criteria defined in certification criteria identified in 45 CFR Part 170 Subpart C of the Health Information Technology: 2015 Edition Health Information Technology (Health IT) Certification Criteria.

Results of the study indicated that the Chirp EHR system was quite satisfactory with regards to effectiveness and efficiency and that the participants were satisfied with the system.

Introduction

The Electronic Health Record System Under Test (EHRUT) tested for this study, the Chirp EHR (V1.0), was specifically designed to present medical information to healthcare providers on desktop computers in standard healthcare settings. This study tested and validated the usability of the Chirp EHR software user interface and provides evidence of the usability of the Chirp EHR with representative exercises and in realistic user conditions. To this end, measures of effectiveness and efficiency, such as time on task, number of errors made, and completion rates were captured during usability testing. Satisfaction was assessed, and user comments collected using two industry-standard questionnaires: The System Usability Scale (SUS) and the Computer System Usability Questionnaire (CSUQ).

Method

Participants

Thirteen (13) individuals (9 women and 4 men) participated in the EHRUT(s) using the Chirp EHR. Participants were physicians, nurses, and/or other healthcare/Health IT practitioners. Participants were recruited from a database of potential participants maintained by The Usability People, LLC. The contacts contained within this database were generated from past participants, and via potential participants responses to postings in Internet and social media sites, and a link at the bottom of The Usability People website. Those who responded to the invitation to take part in the study were directed to an online questionnaire that served as the participant screener. (The screening questionnaire is provided as Appendix A.) Participants meeting the criteria for participation in the study were contacted and scheduled via telephone and email.

Participants in the usability test of the Chirp EHR had a variety of healthcare backgrounds and demographic characteristics.

Table 1 presents participant characteristics, including demographics, professional experience, computing experience, and previous EHR experience. Participant characteristics reflect the audience of current and future users and meet the criteria designated in the 2015 Edition Certification Companion Guide for Safety-enhanced design - 45 CFR 170.315(g)(3). None of the participants were from the vendor organization (Iora Health) that produced and supplied the evaluated system nor did any participant have any direct connection to the testing organization (The Usability People, LLC). All participants were compensated for their time.

Table 1. Participant Characteristics

Part ID	Gender	Age	Education	Role/Title	Professional Experience (Months)	Computer Experience (Months)	Experience with Chirp (Months)	Assistive Tech Needs
P01	Female	40 to 49	Bachelor's degree	Registered Nurse	132	120	0	None
P02	Male	40 to 49	Master's degree	AVP, Regional Operations	60	60	0	None
P03	Male	40 to 49	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Lead Physician	48	48	0	None
P04	Male	50 to 59	Doctorate degree (e.g., MD, DNP, DMD, PhD)	M.D.	300	180	0	None
P05	Female	60 to 69	Master's degree	RN	480	96	0	None
P06	Female	40 to 49	Associate degree	IT Analyst	108	96	0	None
P07	Female	30 to 39	Bachelor's degree	CHR Analyst and Support Specialist	120	84	0	None
P08	Female	60 to 69	Bachelor's degree	RN, BSN, CCRC	480	240	0	None
P09	Female	50 to 59	Associate degree	RN	240	192	0	None
P10	Male	30 to 39	Bachelor's degree	Registered Nurse	120	120	0	None
P11	Female	50 to 59	Associate degree	CT Technologist	264	228	0	None
P12	Female	30 to 39	Bachelor's degree	Registered nurse	24	48	0	None
P13	Female	50 to 59	Master's degree	RN Manager	312	180	0	None

Summary of Participant Characteristics:

Participants had experience with the occupation and expertise that aligns with the capability under testing. The cohort of users who are selected as participants was varied with the product and its intended users and was not limited to clinicians. The demographic characteristics of the test participant characteristics reflected the audience of current and future users.

Gender	
Male	4
Female	9
Age Range	
20 to 29	0
30 to 39	3
40 to 49	4
50 to 59	4
60 to 69	2
70 to 79	0
Education	
Some college credit, no degree	0
Trade technical vocational training	0
Associate degree	3
Bachelor's degree	5
Master's degree	3
Doctorate Degree	2
Years of Experience with Chirp	
None	13
Up to 3 years	0
3 to 5 years	0
5 to 10 years	0
More than 10 years	0

Study Design

The overall objective of this usability test was to uncover areas where the Chirp EHR system performed well – that is, effectively, efficiently, and with satisfaction – and areas where the system failed to serve the clinical documentation and workflow needs of users. Data from this test may be used as a baseline for future tests of updated versions of Chirp and/or for comparing Chirp software with other EHRs presenting the same tasks. In short, this testing serves as both a means to record or benchmark current usability and to identify areas where improvements must be made.

Participants had a range of experience with EHRs in general, but none had any prior experience with all of the Chirp system. Participants completed the Chirp usability study during individual 30-45-minute remote video conference sessions. During the test, each participant interacted with various components of the Chirp software system. Each participant was provided with the same instructions.

The Chirp EHR was evaluated for effectiveness, efficiency and satisfaction as defined by the following measures collected and analyzed for each participant:

- Number of tasks successfully completed without assistance
- Time to complete the tasks
- Number and description of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Tasks

A total of thirteen (13) tasks were constructed by The Usability People, LLC. (in collaboration with Iora Health staff) to be realistic and representative of the activities a user might engage in while using Chirp in actual medical settings. The thirteen (13) tasks were created based upon the criteria specified within the test procedure structure for evaluating conformance of Electronic Health Record (EHR) technology to the certification criteria as defined in 45 CFR Part 170 Subpart C of the Health Information Technology: Standards, Implementation Specifications, and Certification Criteria for Electronic Health Record Technology. The tasks focused on a subset of the nine (9) 2015 Cures Update Edition certification criteria specified by ONC, specifically:

- Section 170.315(a)(1) Computerized provider order entry – medications
- Section 170.315(a)(4) Drug-drug, drug-allergy interaction checks
- Section 170.315(a)(5) Demographics
- Section 170.315(a)(9) Clinical decision support
- Section 170.315(a)(14) Implantable device list
- Section 170.315(b)(2) Clinical information reconciliation and incorporation

A copy of the tasks presented to participants in the usability test of the Chirp EHR can be found in Appendix C.

Test Location

All participants were tested on the Chirp EHR system during remote conferencing sessions. Each participant was requested in advance to secure a quiet room with minimal distractions and a desktop or laptop computer that could connect to the Internet with a remote video conference system. Although the type of computer, operating system and display resolution of the remote participant system was unknown, the system that was used by the test administrator and controlled by the remote participant was a Dell Inspiron Laptop running the Windows 10 professional operating system at a resolution of 1366x768 pixels. During a given remote video conference session, only the test administrator and participant communicated with one another.

The remote usability test session was conducted by a test administrator from the testing organization (*The Usability People, LLC*) working at The Usability People's Fairfax, VA location. A data logger from the testing organization also took detailed notes on each session, including user comments and other ratings following each task. During a session both the test administrator and the data logger could see the participant's screen and hear the participant's comments, questions, and responses.

Test Environment

While the EHRUT typically would be used in a healthcare office, or ambulatory center facility, testing of the Chirp EHR system was conducted via remote connection during individual remote video conference sessions. Each participant connected into a remote video conference session and was connected by the test administrator to the application.

The Chirp system itself ran within a browser on a Windows™ client platform on a LAN connection using a sample database that was set up specifically for the test. Participants used a mouse and keyboard when interacting with the EHRUT and were given remote control of the administrator's workstation to perform the tasks.

Test Forms and Tools

As part of the usability test, several documents and instruments were used. Examples of the documents used during the usability test, including an informed consent form, the tasks, and post-test questionnaires, can be found in Appendices B to E, respectively.

Participant Instructions

The administrator read the following instructions aloud to each participant:

Thank you for participating in this study. Your input is very important. Our session today will last about 45 minutes. During that time, you will use an instance of an electronic health record. I will ask you to a few tasks using this system and answer some questions.

Please note that we are not testing you; we are testing the system. Therefore, if you have any difficulty this may mean that something needs to be improved in the system. I will be here in case you need specific help, but I am not able to instruct you or provide help in how to use the application.

Overall, 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. I did not have any involvement in its creation, so please be honest with your opinions. All of the information that you provide will be kept confidential and your name will not be associated with your comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.

Participants were then given tasks to complete.

Procedure

Upon connection to the online meeting tool, each participant was greeted, his or her identity verified, and matched to a name on the participant schedule. Participant demographic information was also verified, including information regarding overall computer experience and participant experience with the EHUT. Participant names were replaced with participant IDs so that a given individual's data cannot be linked to his/her identity. Prior to beginning testing, each participant reviewed and signed an informed consent form (See Appendix B).

Staff members of the Usability People, a usability test administrator, administered the test. The administrator moderated the session by providing both verbal and written instructions for the overall usability test and for each of the tasks comprising the test. The administrator also monitored task success, path deviations, number and description of errors, and participant verbal comments. A data logger logged task times, obtained post-task rating data, and took notes on participant comments and administrator feedback.

For each of the thirteen (13) tasks, participants were provided written instructions to their computers. Following the administrator's instructions, each participant performed each task by first reading the task then stating in his or her own words their interpretation of the task requirements. When the participant's interpretation matched the actual goal of the task, the administrator instructed the participant to begin and task timing began. Task time was stopped and recorded when the test administrator observed on their workstation that the participant had successfully completed the task. If a participant failed to complete a task before the expected amount of time for each task, that task was marked as "Timed Out." After each task, the test administrator asked the participant, "On a scale from 1 to 5, where 1 is 'Very Difficult and 5 is 'Very Easy,' how satisfied were you with the ease of use for this task?" This same procedure was conducted for each of the thirteen (13) tasks.

Following completion of thirteen (13) EHR tasks, the administrator electronically presented to the participant two post-test questionnaires (The System Usability Scale (SUS), see Appendix D and the Computer System Usability Questionnaire (CSUQ), see Appendix E). After the participant completed both questionnaires, the administrator thanked each participant for his or her time and allowed the participant to make any comments on or ask any questions about the system and/or the tasks presented.

For each session, the participant's schedule, demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were digitally recorded. The system was then reset to proper test conditions for the next participant.

Usability Metrics

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records* (NIST IR 7741, November 2010) 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:

- Effectiveness of the Chirp EHR software by measuring participant success rates and errors.
- Efficiency of the Chirp EHR software Software by measuring the average task time and path deviations.
- Satisfaction with the Chirp EHR software by measuring ease-of-use ratings.

Data Scoring

Table 2 details how tasks were scored, errors evaluated, and the time data analyzed.

Table 2. Scoring Protocols for Effectiveness, Efficiency, and Satisfaction

Measures	Rationale and Scoring
Effectiveness: • Task Success	A task was counted as "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 was calculated for each task and then divided by the total number of times that task was attempted. Results are provided as a percentage.
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 "Fail." No task times were taken for failed attempts. The total number of errors was calculated for each task and divided by the total number of times that task was attempted. Results are presented as the average error rate. Note: Not all deviations are counted as errors
Effectiveness: • Prompted Successes	Because some tasks are dependent upon the successful completion of previous tasks, participants may receive a limited number of "prompts" to help prepare the system data for the pre-requisites for subsequent tasks. When a participant was able to complete the data entry on a task with 3 or fewer prompts, the task was counted as an "Assisted" completion. No task times were recorded for Assisted completions.
Efficiency: • Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if for example, the participant navigated to an incorrect screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control.
Efficiency: • Task Time	Each task was timed from the administrator's prompt "Begin" until said, "Done." If the participant failed to say, "Done," timing 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.
Satisfaction:	

- Ease of Use ratings
- System Satisfaction

Participant's subjective impression of the ease of use of the application was measured by administering both a single post-task question as well as two post-session questionnaires.
After each task, the participant determined on a scale of 1 to 5 their subjective satisfaction with performance on the task. These data are averaged across participants.

To measure participants' confidence in and likeability of the EHR overall, the testing team administered electronic versions of the System Usability Scale (SUS) and the Computer System Usability Questionnaire (CSUQ). See the SUS questionnaire as Appendix D., and the CSUQ as Appendix E.

Results

Data Analysis and Reporting

The results of the usability test of the Chirp EHR software system were analyzed according to the methods described in the Usability Metrics section above and are detailed below. Note that the results should be evaluated relative to the study objectives and goals, as outlined in the study design section above. The data should yield actionable results that, if corrected, yield material, positive impact on user performance.

Reliability

During the data collection phase, it was observed that the system provided a consistent and reliable interface to each participant as they completed their tasks. As each participant completed their assigned tasks, the system provided the same information and responded to their input with the same verbiage and using the same mode of communication (e.g. Pop-up message, or embedded assistance).

Effectiveness and Efficiency

Table 3 presents a summary of overall task performance showing task, mean time on task, task completion rates, mean path deviations and mean task satisfaction:

Table 3. Usability Test Results

Task	Mean Task Time	SD	Completion Rate (%)	Mean # Path Deviations	SD	Mean Task Satisfaction	SD
Task 1- CCDA Reconciliation and Incorporation	5:09	1:11	80%	2.30	2.45	3.50	0.81
Task 2 - Implantable Device List	0:57	0:16	100%	0.30	0.46	4.50	0.92
Task 3 - Implantable Device List - Change	0:43	0:20	100%	0.70	0.90	4.30	0.78
Task 4 - Record medication	1:42	0:23	100%	0.50	0.67	4.60	0.49
Task 5 - Change medication	0:57	0:58	90%	0.60	1.02	4.70	0.64
Task 6A - CPOE medication	1:09	0:33	100%	0.70	0.78	4.30	0.64
Task 6B - View CDS Information	1:05	0:14	100%	1.00	0.00	4.30	0.64
Task 7 - Drug-drug, drug-allergy interaction	0:45	0:15	100%	0.00	0.00	4.40	0.64
Task 8 - Drug-drug, drug-allergy interaction	0:44	0:19	100%	0.20	0.40	4.80	0.40
Task 9 - CCDA Export reconciled data	0:52	0:11	100%	0.80	0.75	4.40	0.66
Task 10 - Demographics and Contact Information	1:13	0:26	100%	0.30	0.64	4.60	0.49
Task 11 - Demographics and Contact Information-Change	1:04	0:12	100%	0.10	0.30	4.60	0.92
Task 12 - Add CDS Intervention	0:15	0:05	100%	0.00	0.00	4.90	0.30
Task 13 - Add Additional intervention	0:50	1:00	90%	1.00	1.48	4.50	0.92

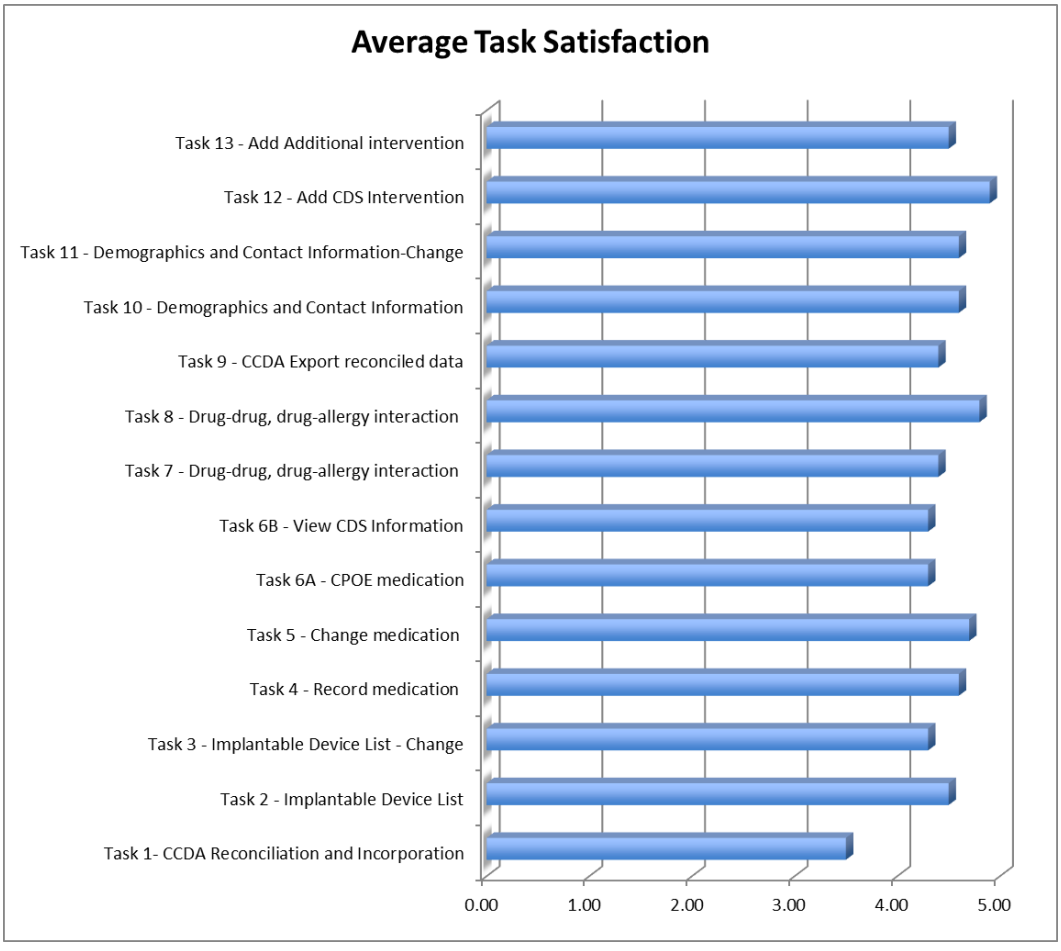
As Table 3 shows, relative to optimal performance standards as defined by Iora Health and The Usability People, participant performance in the Chirp EHR software usability test was quite satisfactory. The overall average task completion rate was ninety-seven (97) percent.

Satisfaction

Individual Task Satisfaction

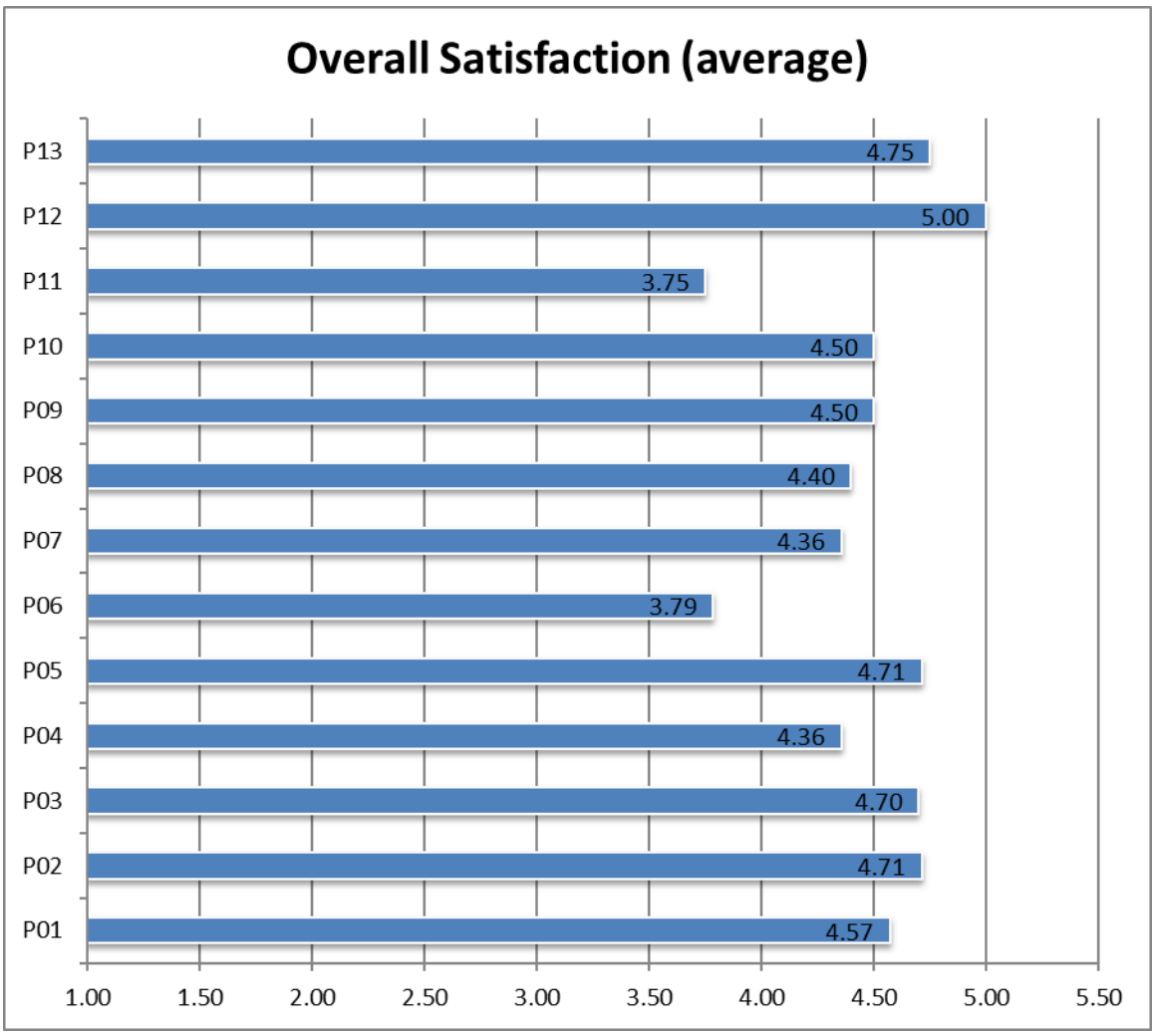
Participants verbally indicated their satisfaction with the ease of use for each task using a scale of "1" ("Very Difficult") to "5," ("Very Easy"). As Figure 1 shows individual task satisfaction ranged from a low 3.5 out of 5 on Task 1 *(CCDA Reconciliation and incorporation)* to a high of 4.9 out 0f 5 on Tasks 12 (Add CDS Intervention*)*.

Figure 1. Satisfaction Ratings of Individual Tasks



Individual Participant Satisfaction

In general, the participants were satisfied with the ease of use of the Chirp EHR software system. The following chart displays overall satisfaction for each participant:



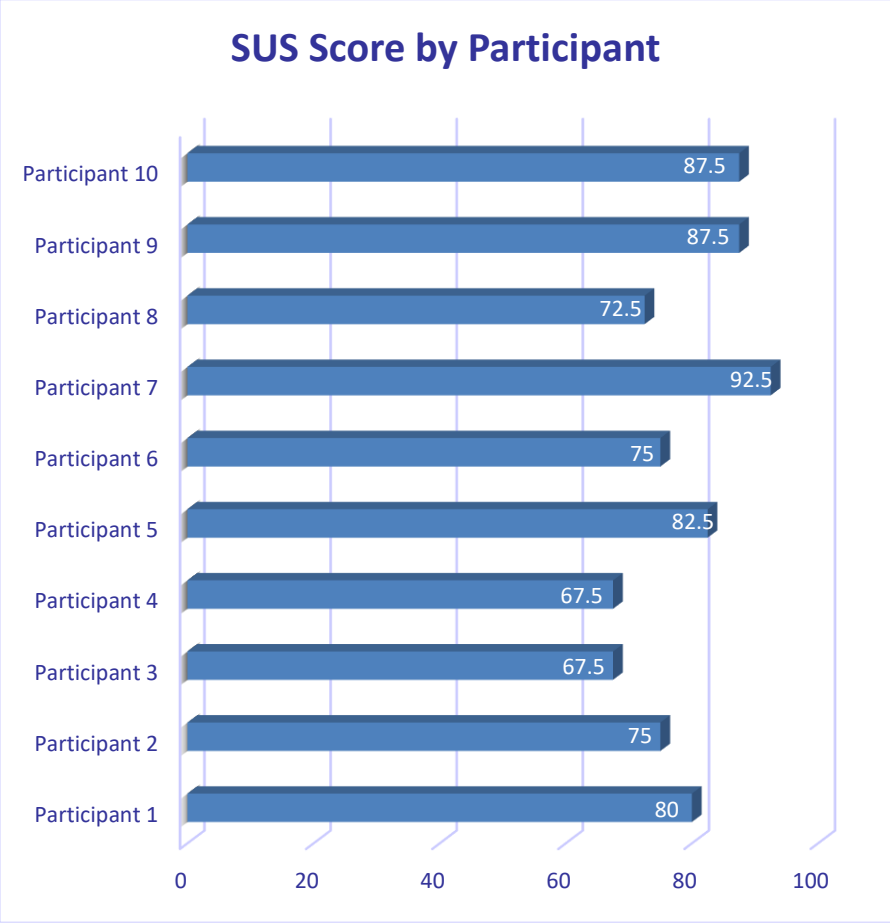
The average overall task satisfaction rate 4.47 out of 5.

System Usability Scale

The System Usability Scale (SUS) is a simple, 10-item Likert-type attitude scale providing a global subjective assessment of usability from the user's perspective (John Brooke at Digital Equipment Company developed the SUS in 1986). The SUS scale is scored from 0 to 100; scores under 60 represent systems with less than optimal usability, scores over 80 are considered better than average. See Appendix D for a copy of the SUS.

The mean SUS score of the Chirp EHR was calculated using ten participants. Overall, participant-users rated their satisfaction with the Chirp software system to be a usable and satisfying EHR.

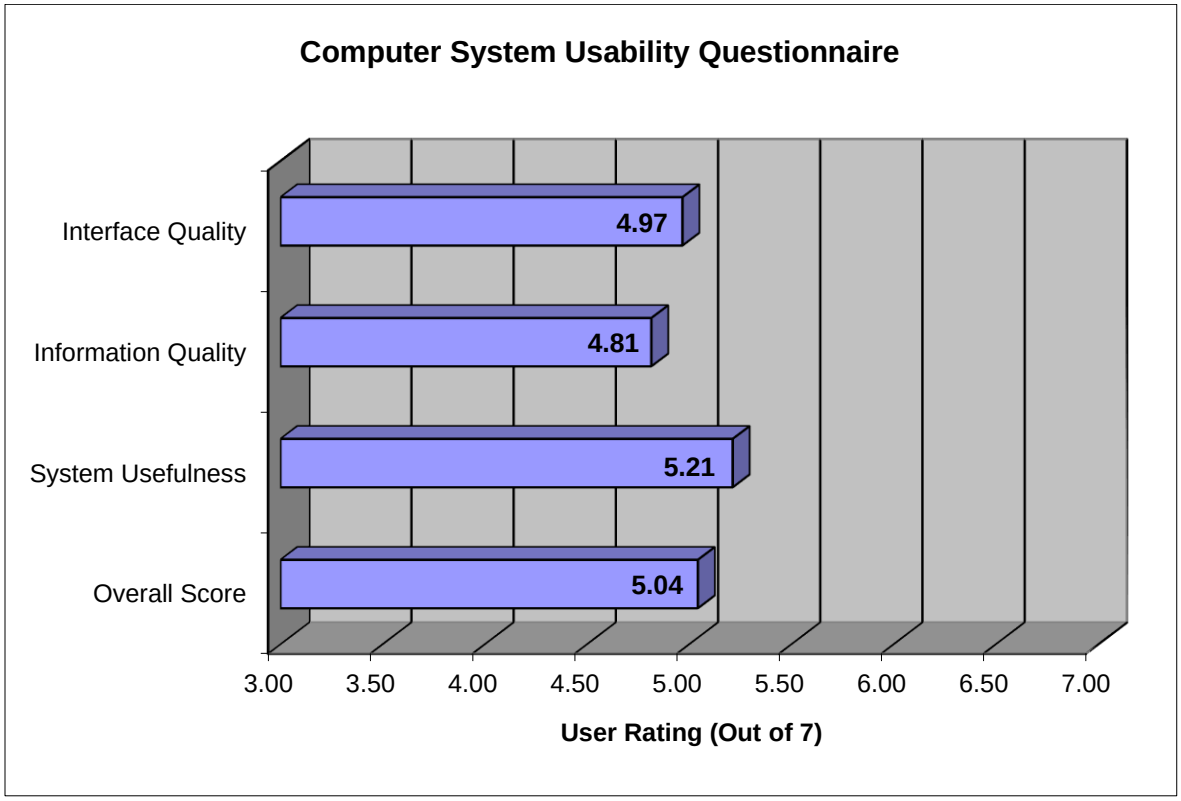
The following chart shows the SUS score by each participant:



Computer System Usability Questionnaire

Using the Computer System Usability Questionnaire (CSUQ; Lewis, J. R. (1995). (See: IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. International Journal of Human-Computer Interaction, 7:1, 57-78).), participants rated each of 19 items of the CSUQ questionnaire on a scale from 1 to 7, with a rating of 7 being most in agreement with the positively-worded item. Responses for each item were summed and averaged to four scales – Interface Quality, Information Quality, System Usefulness- and an overall scale. See Appendix E for a copy of the CSUQ.

Figure 2 displays CUSQ ratings for ten participants using the four scales. In general, participants in the Chirp study rated system usability to be high. On Interface Quality the average score for the participants was 4.97/7; on Information Quality the average score 4.81/7; on System Usefulness the average score was 5.21/7; and the overall average CUSQ score was 5.04/7.



Specific Task Result Details

Participant Number	Task 1 (b2) CCDA Reconciliation and Incorporation			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	4:18	Success	3	3
P02	5:02	Success	0	3
P03	5:32	Success	1	5
P04	6:36	Success	1	3
P05	7:18	Fail	7	3
P06	5:14	Success	0	5
P07	5:41	Success	0	3
P08	3:01	Fail	6	3
P09	3:56	Success	1	4
P10	4:50	Success	4	3

Expected Time on Task	5:00	(SD)
Average Time on Task	5:09	1:11
Average Task Satisfaction	3.50	0.81
Average #Path Deviations	2.30	2.45
Percent Success	80%	

Participant Number	Task 2 (a14) Implantable Device List			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	1:15	Success	0	5
P02	0:31	Success	0	5
P03	0:39	Success	0	5
P04	1:15	Success	1	5
P05	1:02	Success	0	5
P06	1:09	Success	0	2
P07	1:16	Success	0	5
P08	0:47	Success	1	4
P09	0:45	Success	0	5
P10	0:55	Success	1	4

Expected Time on Task	0:30	(SD)
Average Time on Task	0:57	0:16
Average Task Satisfaction	4.50	0.92
Average #Path Deviations	0.30	0.46
Percent Success	100%	

Participant Number	Task 3 (a14) Implantable Device List – Change			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	0:38	Success	1	5
P02	0:32	Success	0	5
P03	0:41	Success	0	5
P04	0:39	Success	0	3
P05	1:01	Success	0	5
P06	0:39	Success	0	3
P07	0:24	Success	1	4
P08	1:35	Success	3	4
P09	0:38	Success	1	4
P10	0:27	Success	1	5

Expected Time on Task	0:30	(SD)
Average Time on Task	0:43	0:20
Average Task Satisfaction	4.30	0.78
Average #Path Deviations	0.70	0.90
Percent Success	100%	

Participant Number	Task 4 (a1) Record medication			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	1:10	Success	0	5
P02	1:28	Success	0	5
P03	1:30	Success	0	5
P04	1:44	Success	0	4
P05	2:28	Success	1	5
P06	1:14	Success	0	4
P07	1:42	Success	1	5
P08	1:41	Success	0	5
P09	1:54	Success	1	4
P10	2:11	Success	2	4

Expected Time on Task	1:30	(SD)
Average Time on Task	1:42	0:23
Average Task Satisfaction	4.60	0.49
Average #Path Deviations	0.50	0.67
Percent Success	100%	

Participant Number	Task 5 (a1) Change medication			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	0:39	Success	0	5
P02	1:19	Success	0	5
P03	0:27	Success	0	5
P04	0:36	Success	0	5
P05	3:39	Timeout	3	5
P06	0:26	Success	0	3
P07	2:21	Success	1	4
P08	0:56	Success	2	5
P09	0:48	Success	0	5
P10	1:01	Success	0	5

Expected Time on Task	1:00	(SD)
Average Time on Task	0:57	0:58
Average Task Satisfaction	4.70	0.64
Average #Path Deviations	0.60	1.02
Percent Success	90%	

Participant Number	Task 6A - CPOE medication			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	1:46	Success	0	4
P02	0:35	Success	2	5
P03	2:14	Success	0	4
P04	0:30	Success	0	4
P05	0:44	Success	1	5
P06	1:19	Success	1	3
P07	1:13	Success	1	4
P08	1:22	Success	0	5
P09	1:14	Success	2	4
P10	0:29	Success	0	5

Expected Time on Task	1:30	(SD)
Average Time on Task	1:09	0:33
Average Task Satisfaction	4.30	0.64
Average #Path Deviations	0.70	0.78
Percent Success	100%	

Participant Number	Task 6B - View CDS Information			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	1:18	Success	0	4
P02	0:54	Success	0	5
P03	1:03	Success	0	4
P04	1:31	Success	0	4
P05	1:26	Success	0	5
P06	0:58	Success	0	3
P07	0:59	Success	0	4
P08	0:52	Success	0	5
P09	0:54	Success	0	4
P10	0:55	Success	0	5

Expected Time on Task	1:30	(SD)
Average Time on Task	1:05	0:14
Average Task Satisfaction	4.30	0.64
Average #Path Deviations	0.00	0.00
Percent Success	100%	

Participant Number	Task 7 - Drug-drug, drug-allergy interaction			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	0:35	Success	0	5
P02	0:42	Success	0	4
P03	0:37	Success	0	4
P04	0:53	Success	0	4
P05	1:15	Success	0	4
P06	0:43	Success	0	4
P07	0:33	Success	0	4
P08	1:08	Success	0	5
P09	0:32	Success	0	5
P10	0:29	Success	0	5

Expected Time on Task	1:00	(SD)
Average Time on Task	0:45	0:15
Average Task Satisfaction	4.40	0.49
Average #Path Deviations	0.00	0.00
Percent Success	100%	

Participant Number	Task 8 - Drug-drug, drug-allergy interaction			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	0:27	Success	0	5
P02	0:31	Success	0	5
P03	0:39	Success	0	5
P04	1:02	Success	0	5
P05	1:23	Success	1	5
P06	0:58	Success	0	4
P07	0:31	Success	0	5
P08	0:58	Success	1	4
P09	0:29	Success	0	5
P10	0:20	Success	0	5

Expected Time on Task	0:45	(SD)
Average Time on Task	0:44	0:19
Average Task Satisfaction	4.80	0.40
Average #Path Deviations	0.20	0.40
Percent Success	100%	

Participant Number	Task 9 - CCDA Export reconciled data			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	0:47	Success	2	4
P02	0:58	Success	0	4
P03	0:48	Success	0	5
P04	0:38	Success	1	5
P05	1:08	Success	1	5
P06	0:43	Success	0	5
P07	0:49	Success	0	5
P08	1:14	Success	1	4
P09	0:40	Success	1	3
P10	0:58	Success	2	4

Expected Time on Task	1:00	(SD)
Average Time on Task	0:52	0:11
Average Task Satisfaction	4.4	0.66
Average #Path Deviations	0.80	0.75
Percent Success	100%	

Participant Number	Task 10 - Demographics and Contact Information			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	1:54	Success	1	4
P02	0:50	Success	0	5
P04	0:41	Success	0	5
P05	2:02	Success	2	4
P06	1:03	Success	0	5
P07	1:13	Success	0	5
P09	1:01	Success	0	4
P11	1:25	Success	0	4
P12	1:13	Success	0	5
P13	0:46	Success	0	5

Expected Time on Task	1:00	(SD)
Average Time on Task	1:13	0:26
Average Task Satisfaction	4.6	0.49
Average #Path Deviations	0.30	0.64
Percent Success	100%	

Participant Number	Task 11 - Demographics and Contact Information-Change			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	1:15	Success	0	5
P02	0:58	Success	0	5
P04	0:39	Success	0	5
P05	1:15	Success	1	5
P06	1:02	Success	0	5
P07	1:09	Success	0	2
P09	1:16	Success	0	5
P11	0:47	Success	0	5
P12	1:13	Success	0	5
P13	1:03	Success	0	4

Expected Time on Task	1:00	(SD)
Average Time on Task	1:04	0:12
Average Task Satisfaction	4.6	0.92
Average #Path Deviations	0.10	0.30
Percent Success	100%	

Participant Number	Task 12 - Add CDS Intervention			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	0:13	Success	0	5
P02	0:09	Success	0	5
P04	0:08	Success	0	5
P05	0:27	Success	0	5
P06	0:11	Success	0	5
P07	0:18	Success	0	5
P09	0:18	Success	0	5
P11	0:16	Success	0	4
P12	0:10	Success	0	5
P13	0:17	Success	0	5

Expected Time on Task	1:00	(SD)
Average Time on Task	0:15	0:05
Average Task Satisfaction	4.9	0.30
Average #Path Deviations	0.00	0.00
Percent Success	100%	

Participant Number	Task 13 - Add Additional intervention			
	Task Time	Outcome	# Path Deviations	Task Satisfaction
P01	0:24	Success	0	5
P02	0:25	Success	0	5
P04	0:51	Success	1	5
P05	1:33	Success	2	4
P06	0:55	Success	1	5
P07	1:13	Success	1	4
P09	0:52	Success	0	5
P11	4:00	Fail	5	2
P12	0:35	Success	0	5
P13	0:45	Success	0	5

Expected Time on Task	0:30	(SD)
Average Time on Task	0:50	1:00
Average Task Satisfaction	4.5	0.92
Average #Path Deviations	1.00	1.48
Percent Success	90%	

Discussion of Findings

In general, the participants performed very well and were satisfied with Chirp EHR system. Some participants struggled with some portions of a few tasks but in general most were able to successfully complete a large majority of the tasks with little or no difficulty. Given that no participants had any prior experience with the system, this shows that the Chirp EHR is an easy-to-use and an easy-to-learn EHR. Participants were mostly able to perform all tasks successfully on their own with no assistance or external documentation. The participant average performance rate was high. The Chirp system appears to be a usable EHR.

Effectiveness

Of the thirteen (13) tasks presented, a large majority of the tasks were successfully completed by all of the participants. Over all of participants, the mean successful task competition rate was high with an overall average rate of ninety-seven (97) percent indicating that in general the participants had little or no difficulty completing the tasks.

The amount of prior experience with EHR systems was related to successful task performance and error rates; participants with more prior experience were more slightly likely to successfully complete tasks with less errors than those with less prior experience.

Efficiency

Participants who successfully completed tasks generally completed those tasks within an acceptable time. Some tasks were completed more quickly than the calculated optimal time, while several tasks took slightly longer than expected. The tasks that took the longest required the participants to navigate to an unfamiliar portion of a page, interact with a workflow, locate and select specific actions. Those tasks may be performed more quickly with a minor set of updates to the user interface and/or the user experience.

Some participants made errors when attempting to navigate toward solving their assigned tasks. These errors may be associated with those participants not being familiar with features and

not understanding the presented workflow of the Chirp software system. As noted above, prior experience with EHR systems was related to successful task completion.

Satisfaction

Participants were satisfied with the Chirp Software system; ratings on the SUS (mean = 79 out of a possible 100) and the CSUQ (Overall score = 5.04 out of a possible 7.0) demonstrated a high degree of satisfaction with the system.

On the CSUQ, participants ranked the scale "(System Usefulness)" highest of the three scales, suggesting that users felt that the system would likely solve their tasks in an effective and efficient manner. Individual task satisfaction ratings were related to individual user performance. Those participants who were able to successfully complete tasks were also more likely to rank those tasks as satisfying, while those participants who did poorly or were not able to complete a task ranked those tasks as less satisfying.

Summary of Major Findings

This evaluation demonstrated that the Chirp EHR system is a usable system with a short learning curve. Participants with no experience using the system experienced very little difficulty understanding the navigation and information architecture. Participants with more overall EHR experience were slightly more able to solve most tasks without difficulty or error.

Risk Analysis

The following table presents a list of tasks presented with the risk of error as observed during the testing.

Table 5. Risk Analysis

Task	Description	Percent Complete	Risk Status
Task 1- CCDA Reconciliation and Incorporation	Order patient medications	80%	Low
Task 2 - Implantable Device List	Enter Implantable device information	100%	None
Task 3 - Implantable Device List - Change	Modify Implantable device information	100%	None
Task 4 - Record medication	Record a medication	100%	None
Task 5 - Change medication	Modify a medication	90%	Low
Task 6A - CPOE medication	Access CPOE Medication	100%	None
Task 6B - View CDS Information	View Decision support info for medication	100%	None
Task 7 - Drug-drug, drug-allergy interaction	View CDS drug information	100%	None
Task 8 - Drug-drug, drug-allergy interaction	View CDS allergy information	100%	None
Task 9 - CCDA Export reconciled data	Export CCDA data	100%	None
Task 10 - Demographics and Contact Information	Enter Demographic Information	100%	None
Task 11 - Demographics and Contact Information-Change	Modify Demographic Information	100%	None
Task 12 - Add CDS Intervention	Add Clinical Decision Support Intervention	100%	None
Task 13 - Add Additional intervention	Add Additional CDS Intervention	90%	Low

Areas for Improvement

The following is a partial list of potential areas for improvement.

Making these and other minor enhancements will improve the overall user experience of the Chirp system and increase the effectiveness, efficiency, and satisfaction for both experienced and novice users.

- **Indication of Required Fields**
    - A frequent error was caused when participants attempted to submit a form within the interface that contained required fields that were not completed. This is likely because the system did not provide a clear indication of which fields are required. Adding a consistent visual indication of required fields would likely eliminate many of these errors observed.
- **Pop-up dialogs closed when the user clicked outside of the dialog box**
    - Many errors occurred when users attempted to interact with a pop-up dialog. The user tried to scroll to the bottom of these dialogs, and the dialog would close if they clicked on the incorrect scroll bar or on the background. Modal Dialogs should remain open until the user interacts with the controls associated with the dialog.
- **Inconsistent wording across User Interface**
    - Several participants commented that sometimes the interface presented a feature or function that did not exactly match their expectations. For example, when participants clicked on something that said “save” they expected that that item is saved, not “staged”. Creating a textual information that is consistent across all functional areas would help to improve the overall user experience of the system.

Appendices

Appendix A: Recruiting Screener

1. Are you male or female?

2. Have you participated in a focus group or usability test in the past *6* months?

3. Do you, or does anyone in your home work in marketing research, usability research, and/or web design?

4. Do you, or does anyone in your home, have a commercial or research interest in an electronic health record software or consulting company?

5. Which of the following best describes your age?

    20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-99, 100 and older.

6. Which of the following best describes your education level?
- No high school diploma
- High school graduate, diploma or the equivalent
- Some college credit, no degree
- Trade technical vocational training
- Associate degree
- Bachelor's degree
- Master's degree
- Doctorate degree (e.g., MD, DNP, DMD, PhD)

7. Do you require any assistive technologies to use a computer?

8. Please describe your medical or nursing credentials

9. What is your current job title?

10. How long have you held this position? (number of years):

11. What type of facility do you work in and what is your role there?

12. How are medical records handled at your (main) workplace?

    _____All Paper _____Some Paper/Some Electronic ___All Electronic

13. How many EHRs do you use or have you worked with?

14. How many years have you used an electronic health record?

15. How many years have you used the Chirp system?

16. How many years have you used computers?

17 About how many hours per week do you spend using a computer?

18. What computer platform(s) do you usually use?

19. In the last month, about how often have you used an electronic health record?

_____Did not use last month ___Every day _____A few times a week.

Appendix B: Informed Consent Form

The Usability People would like to thank you for participating in this study. The purpose of this study is to evaluate an electronic health records system. If you decide to participate, you will be asked to perform several tasks using the prototype and give your feedback. The study will last about *60* minutes.

Agreement

I understand and agree that as a voluntary participant in the present study conducted by The Usability People. I am free to withdraw consent or discontinue participation at any time. I understand and agree to participate in the study conducted and recorded by The Usability People.

I understand and consent to the use and release of the video recording by The Usability People. I understand that the information and video is for research purposes only and that my name and image will not be used for any purpose other than research. I relinquish any rights to the video and understand the video recording may be copied and used by The Usability People without further permission.

I understand and agree that the purpose of this study is to make software applications more useful and usable in the future.

I understand and agree that the data collected from this study may be shared outside of The Usability People. I understand and agree that data confidentiality is assured, because only de-identified data – i.e., identification numbers not names – will be used in analysis and reporting of the results.

I agree to immediately raise any concerns or areas of discomfort with the study administrator. I understand that I can leave at any time.

Please check one of the following:

____YES, I have read the above statement and agree to be a participant.

____NO, I choose not to participate in this study.

Signature: _____ Date _____

Appendix C: Participant Guide

ORIENTATION and INTRODUCTION

Thank you for participating in this study. The session today will last approximately 30-45 minutes**.** During this session, you will look at a special version of the Chirp EHR. The product you will be using today may not be exactly like the software that you may be used to. Some of the data may not make sense, as it is merely placeholder data.

The purpose of this study is to gauge the *usability* of the software. We are interested in how easy (or how difficult) this version is to use, and learning about which of the features and benefits would be most useful to you. We also want to know how we could improve it.

The moderator will guide you through the testing process. Please use the software as you would in a real clinical setting. You will be asked to save your comments until the end of a task or the end of the session to discuss your thoughts freely.

Here are some things you should know about your participation in this session

- The Moderator will guide you through each task.
- **Please do not work ahead**.
- If you notice an error, just cancel out and continue to the task. If you feel lost or have difficulty completing the scenario, please inform the Moderator.
- The Testing session will be audio recorded the audio for further study.
- All information will be kept confidential. Your name will not be associated with your comments at any time.

Fictitious patient scenarios have been created and pre-loaded in the system software. The Moderator will ask you to complete several tasks using the system. You will be asked to answer some questions and to complete some tasks on your own. Try to complete tasks as quickly as possible, with the fewest possible errors or deviations.

Please do not do anything more than asked.

Your patient NANCY NEWMAN has arrived for their scheduled appointment. She has been referred to our organization by an outside primary care facility

Task 1 Clinical Information Reconciliation and Incorporation

Your patient has recently seen been seen by an outside primary care facility and that provider has sent a message that contains important Clinical information. This information is available as "Nancy Newman Health Exchange Document" in the communications inbox.

You need to reconcile this external information into the Chirp patient chart for **Nancy Newman**

View the appropriate message, reconcile and incorporate the Issues, Medications, and Allergies present in the attached CCDA file from the outside provider with your patient's existing records.

Make sure that you enter "Nancy Newman" as the **PATIENT** so that Chirp knows which patient record to incorporate into and reconcile.

Issues (reconcile and select appropriate ICD-10 code)
Chronic rejection of renal transplant
Essential hypertension
Fever (Do not Accept)
Sever Hypothyroidism

Medications
Aranesp (0.5 MG/ML)
Tylenol 500mg

Allergy
Ampicillin
Penicillin G benzathine

Finalize the reconciliation, view the reconciled list and verbally report that the Issues, Medications, and Allergies have been incorporated into the record.

Task 2 Implantable Device List

When discussing their surgical history, the patient reports that they have an implanted medical device and provided you with a card that lists the following Device ID:

(01)00643169007222(17)160128(21)BLC200461H

(This UDI number should be in your 'clipboard' and a copy has also been saved as a txt file in your desktop).

Please add this device information as a "Pacemaker" into the Surgical History list

View the details of this device and **Verbally state the device description, identifiers, and attributes of this device.**

Task 3 Implantable Device List - Change

Upon further discussion with the patient, they report that the implantable device was actually removed a few years ago.

Please change the device status to an **Inactive status** and set the Inactivity reason to **Removed from Patient.**

Verify that the changes have been saved

Task 4 Record medication

Your patient reports that they are taking **Aspirin** for some pain they have been feeling in their left elbow. View the patient's current medication list and add the following medication:

Aspirin Tablet 325 mg once every 4 hours.

Verify and/or Enter and save this information into the EHR.

Task 5 Change medication

Upon further conversation with the patient, they told you that they are not in fact taking aspirin but have been taking Aleve.
Because the aspirin was **entered by error**, change the medication list so that the actual drug is listed in the EHR.

Aleve (Naproxen sodium) Capsule 220 once every 6 to 8 hours

Verify and/or Enter and save this information into the EHR.

Verify that the changes have been saved and are displayed in the medication list.

Task 6 CPOE medication

A)

You noticed from the information that was reconciled from the outside provider that your patient has been diagnosed with high blood pressure. Create a new prescription for your patient to help treat their hypertension.

Prescribe **Lisinopril** (tablet) 10 mg, once a day for 30 days

B)

Notice the clinical alert associated with the Lisinopril. Expand the alert to see the clinical reference associated with the alert and view the details of the reference

You decided to continue with this medication and prescribe it anyway. Review and then save and add the RX to the patient record.

Verify that the Lisinopril has been added to the EHR by examining the pending RX list

Task 7 Drug-drug interaction checks

You noticed from the information received from the outside provider that your patient has been diagnosed with high blood pressure. You want to add a new drug to help treat their hypertension

Prescribe **Benazepril (tablet)** 10 mg, once a day for 30 days

The system will display warning messages regarding this new medication.

Review the warning details and set the system to no longer show the interaction warning for this patient

DO NOT continue with the selection.

Task 8 Drug-allergy interaction

The patient presents with infected sores on their left forearm. Add the following medication:

Penicillin V potassium tablet 250 mg 2 times per day for the infection.

The system will display a warning message regarding this new medication.

Review the warning message(s) and then **DO NOT** continue with the order.

Task 9 CCDA View reconciled data

The patient is done with their appointment. The patient goes to the checkout desk. While making their next appointment, the patient asks to have a fully up-to-date readout as a result of today's visit.

Use the **Export Record** function to generate an export of the Patient Record.
(The system will automatically create a task for this export)

View the Task lists and verify that the Patient Record Export has been completed.

Your patient **____ Newman** has completed their scheduled appointment, but before they leave, they want to make sure that their demographic information is up-to-date and correct.

Task 10 (a.5) Demographics and Contact Information

Before ending their appointment, you need to verify or enter some of the patient's demographic information, including identity information, and contact details that are stored by the EHR
View and edit/enter the following details:

Identity Details:
Sex Assigned at Birth ***Female***
Gender Identity '**Identifies as Woman/Female'**
Sexual Orientation '**Straight or Heterosexual'**

Ethnicity ' '**Not Hispanic or Latino**'

Racial Background '**Black or African American'**

Save the patient identity details into the EHR

Contact Details*:*
Preferred **Language** '**Prefer not to say'**

Save the patient contact details into the EHR

Make sure that any changes have been saved.

Task 11 (a.5) Demographics and Contact Information - Changes

After additional consultation with the patient, they ask you to make the following additional changes to their demographic information, including the following identity information, and contact details:
Set the following values:

***Contact Details**:*
***Preferred* Language** **'English'**

Save the patient contact details into the EHR

Identity Details:
Date of Birth ***'01/06/1959'***
Sex Assigned at Birth ***'Female'***
Gender Identity Identifies as **'Non-conforming Gender'**
Sexual Orientation '**Prefer not to Say**

Ethnicity '**Not Hispanic or Latino**'

Patient Race '**Asian'**

Save the patient identity details into the EHR

Verify that these changes have been saved.

Task 12 Add CDS intervention

The Chirp EHR uses a set of "Markers" that are assigned to an automatic clinical decision support system.

When you subscribe to a set of markers, the system will automatically monitor the values associated with these markers and provide valuable clinical information, when necessary, in real-time.

For the current patient record, **Subscribe to the following markers:**
- **'Quality Priorities'**
'Hypertension'
'Care of older adults'

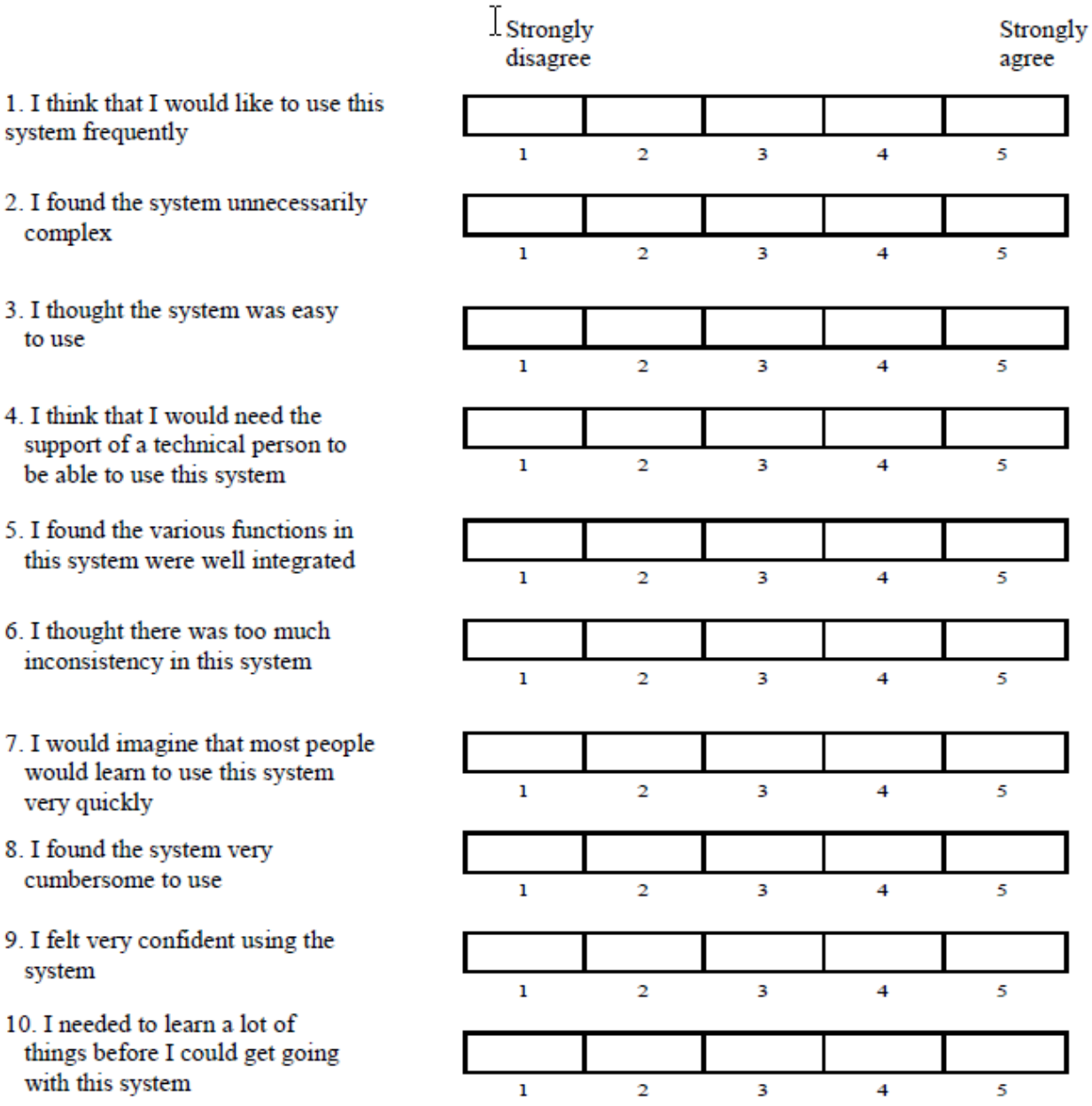
Task 13 Add Additional intervention

During your examination, you discover the possible early signs of COPD in your patient. You'd like to also add this as an issue in Chirp, so that the system will help you monitor any changes in their condition.

Create a new issue
Set the following values
Title: **COPD**
Assign the ICD-10 code of **J44.9**
Assign the marker set to **COPD** to enable automatic monitoring

Save the issue and *verify that the issue has been saved.*

Appendix D: System Usability Scale Questionnaire



Appendix E: Computer System Usability Questionnaire

Please provide your impression of the usability of the system by answering each of the questions below:

1. Overall, I am satisfied with how easy it is to use this system
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

2. It was simple to use this system
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

3. I can effectively complete my work using this system
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

4. I am able to complete my work quickly using this system
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

5. I am able to efficiently complete my work using this system
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

6. I feel comfortable using this system
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

7. It was easy to learn to use this system
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

8. I believe I became productive quickly using this system
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

9. The system gives error messages that clearly tell me how to fix problems
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

10. Whenever I make a mistake using the system, I recover easily and quickly
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree
11. The information (such as online help, on-screen messages, and other documentation) provided with this system is clear
Strongly 1 2 3 4 5 6 7 NA Strongly
Disagree Agree

12. It is easy to find the information I needed

Strongly Disagree 1 2 3 4 5 6 7 NA Strongly Agree

13. The information provided for the system is easy to understand
Strongly Disagree 1 2 3 4 5 6 7 NA Strongly Agree

14. The information is effective in helping me complete the tasks and scenarios
Strongly Disagree 1 2 3 4 5 6 7 NA Strongly Agree

15. The organization of information on the system screens is clear
Strongly Disagree 1 2 3 4 5 6 7 NA Strongly Agree

16. The interface of this system is pleasant
Strongly Disagree 1 2 3 4 5 6 7 NA Strongly Agree

17. I like using the interface of this system
Strongly Disagree 1 2 3 4 5 6 7 NA Strongly Agree

18. This system has all the functions and capabilities I expect it to have
Strongly Disagree 1 2 3 4 5 6 7 NA Strongly Agree

19. Overall, I am satisfied with this system
Strongly Disagree 1 2 3 4 5 6 7 NA Strongly Agree

Appendix F. References

NISTIR 7741 - NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records, Robert M. Schumacher, User Centric, Inc., Svetlana Z. Lowry, Information Access division, Information Technology Laboratory, National Institute of Standards and Technology, U.S. Department of Commerce, National Institute of Standards and Technology, Version 0.2, 15-Nov.2010.

NISTIR 7742 - NIST Customized Common Industry Format Template for Electronic Health Record Usability Testing, Robert M. Schumacher, User Centric, Inc., Svetlana Z. Lowry, Information Access division, Information Technology Laboratory, National Institute of Standards and Technology, U.S. Department of Commerce, National Institute of Standards and Technology, Version 0.2, 15-Nov.2010.

NISTIR 7804 - Technical Evaluation, Testing, and Validation of the Usability of Electronic Health Records, Svetlana Z. Lowry, Matthew T. Quinn, Mala Ramaiah, Robert M. Schumacher, Emily s. Patterson, Robert North, , Information Access division, Information Technology Laboratory, National Institute of Standards and Technology, U.S. Department of Commerce, National Institute of Standards and Technology, Version 0.2, 15-Feb.2012.

ONC - 2015 Edition Certification Companion Guide Safety- enhanced design - 45 CFR 170.315(g)(3b) 3/30/2016

ONC - 2015 Edition Certification - Test Procedure for §170.315(g)(3) Safety Enhanced Design

CSUQ - Lewis, J. R. (1995). IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. International Journal of Human-Computer Interaction, 7, 57-78.

SUS Brooke, J. (1996) SUS: a "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & A. L. McClelland (eds.) Usability Evaluation in Industry. London: Taylor and Francis.

EHR Usability Test Report of Chirp vs 1.1

Decision Support Intervention 315(b)(11) Supplemental Report

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: September 24-October 4, 2024

Date of Report: October 11, 2024

Table of Contents

EXECUTIVE SUMMARY

A usability test of Chirp vs 1.1 was conducted virtually during September 24-October 4, 2024, 2024 by Chart Lux Consulting. The purpose of this test was to test and validate the usability of the enhancements made for the ONC criterion § 170.315(b)(11) Decision Support Intervention functionality. This EHR Under Test (EHRUT) was previously certified and usability tested for the functionality of ONC criterion § 170.315(a)(9). The tasks and test results described in this supplemental report focus on the delta updates for § 170.315(b)(11) compared to its predecessor § 170.315(a)(9). Please refer to original usability test report covering § 170.315(a)(9) tasks for additional information.

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 which cover the changes from § 170.315(a)(9) to § 170.315(b)(11):

- Admin User Selects Evidenced-based DSI and Access / Record / Change Source Attributes
- User Triggers Evidenced-based DSI and Provides User Feedback
- Admin User Exports User Feedback
- Admin User Configures User-supplied Predictive DSI and Records / Changes / Access Source Attributes and Then User Triggers User-supplied Predictive DSI

During the 30 minute one-on-one usability test, each participant was greeted by the administrator, and they were given a instructions for the test event (included in Appendix B). Participants had varied experience with previous versions of this EHRUT, but this specific version was new to them and had some new features they had never experienced before. 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)
Admin User Selects Evidenced-based DSI and Access / Record / Change Source Attributes	10	100% (0%)	10 / 9	62 (14)	14 / 45	0% (0%)	4.7 (0.46)
User Triggers Evidenced-based DSI and Provides User Feedback	10	100% (0%)	9 / 7	38 (19)	19 / 15	10% (30%)	5.0 (0.0)
Admin User Exports User Feedback	10	100% (0%)	3 / 3	15 (0)	0 / 15	0% (0%)	5.0 (0.0)
Admin User Configures User-supplied Predictive DSI and Records / Changes / Access Source Attributes and Then User Triggers User-supplied Predictive DSI	10	100% (0%)	10 / 9	78 (9)	9 / 60	0% (0%)	5.0 (0.0)

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 71.75. 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 and ease of use. Capturing user feedback on DSI was a new concept for them, and they

welcomed the opportunity to supply this. Accessing of the user feedback as an administrative user was determined to be very simple and straightforward.

Areas for Improvement

While results were good and high marks given, some comments were made that system required too much “clicking” and could be made simpler. While the look of the design is very uncluttered, some participants were frustrated they had to enter in data directly when they felt at places it could be automatically done before of a previous action.

Usability Report

Introduction

The EHR Under Test (EHRUT) tested for this study was Chirp version 1.1, designed to test and validate the usability of the enhancements made for the ONC criterion § 170.315(b)(11) Decision Support Intervention functionality. This EHRUT was previously certified and usability tested for the functionality of ONC criterion § 170.315(a)(9). The tasks and test results described in this supplemental report focus on the delta updates for § 170.315(b)(11) compared to its predecessor § 170.315(a)(9). Please refer to original usability test report covering § 170.315(a)(9) tasks for additional information. The usability testing attempted to represent realistic exercises and conditions associated with the § 170.315(b)(11) functionality within the EHRUT.

The purpose of this study was to test and validate the usability of the current user interface and provide evidence of usability in the EHRUT for the associated tasks in this report. 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

Chirp 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 nurse practitioner. Participants were recruited by the developer, and participants had no direct connection to the development of the EHRUT. While the Participants were familiar with older versions of the EHRUT, this specific version was new to them and had some new features they had never experienced before. 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
101	F	30-39	Doctorate	Physician	72	108	36	None
102	M	30-39	Doctorate	Physician	48	108	24	None
103	F	40-49	Doctorate	Physician	168	216	54	None
104	F	40-49	Doctorate	Physician	108	192	48	None
105	F	40-49	Master's Degree	Nurse Practitioner	96	96	36	None
106	F	40-49	Doctorate	Physician	288	288	54	None
107	F	30-39	Doctorate	Physician	96	204	36	None
108	F	30-39	Doctorate	Physician	156	156	60	None
109	F	30-39	Doctorate	Physician	156	156	48	None
110	M	30-39	Doctorate	Physician	96	132	48	None

All ten participants (matching the demographics in the section on Participants) were recruited and participated in the usability test. Participants were scheduled for 30 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. Admin User Selects Evidenced-based DSI and Access / Record / Change Source Attributes (Low Risk)
2. User Triggers Evidenced-based DSI and Provides User Feedback (Medium Risk)
3. Admin User Exports User Feedback (Low Risk)
4. Admin User Configures User-supplied Predictive DSI and Records / Changes / Access Source Attributes and Then User Triggers User-supplied Predictive DSI (Medium Risk)

Procedures

Test participants were scheduled for 30 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 was read the scenario and then provided instructions on the task to perform. Task timing began once the administrator finished reading the question. 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 setup 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 1Life Healthcare 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 all required 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 Chirp by measuring participant success rates and errors
2. Efficiency of Chirp by measuring the average task time and path deviations
3. Satisfaction with Chirp 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 Chirp 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)
Admin User Selects Evidenced-based DSI and Access / Record / Change Source Attributes	10	100% (0%)	10 / 9	62 (14)	14 / 45	0% (0%)	4.7 (0.46)
User Triggers Evidenced-based DSI and Provides User Feedback	10	100% (0%)	9 / 7	38 (19)	19 / 15	10% (30%)	5.0 (0.0)
Admin User Exports User Feedback	10	100% (0%)	3 / 3	15 (0)	0 / 15	0% (0%)	5.0 (0.0)
Admin User Configures User-supplied Predictive DSI and Records / Changes / Access Source Attributes and Then User Triggers User-supplied Predictive DSI	10	100% (0%)	10 / 9	78 (9)	9 / 60	0% (0%)	5.0 (0.0)

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 71.75. 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 switch between opened browser tabs 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 71.75, with the lowest score being 52.5 and the highest score being 87.5.

Major Findings

Participants gave the system high marks and noted it to be very usable and praised its simplicity and ease of use. Capturing user feedback on DSI was a new concept for them, and they welcomed the opportunity to supply this. Accessing of the user feedback as an administrative user was determined to be very simple and straightforward

Areas for Improvement

While results were good and high marks given, some comments were made that system required too much “clicking” and could be made simpler. While the look of the design is very uncluttered, some participants were frustrated they had to enter in data directly when they felt at places it could be automatically done before of a previous action

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: Admin User Selects Evidenced-based DSI and Access / Record / Change Source Attributes

The authorized admin user goes to EHR to select or activate an evidenced-based DSI. After evidenced-based DSI is selected, the admin user will access source attributes and then record and change them.

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 and described above.
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. Results of Pathway Choice
    a. Correct
    b. Minor Deviations/Cycle (describe below)
    c. Major Deviations (describe below)
    d. Comments:
7. Record Errors and Verbalizations:
8. 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(b)(11) Decision Support Intervention

Task 2: User Triggers Evidenced-based DSI and Provides User Feedback

The user engages the previously activated evidenced-based DSI and triggers an intervention. User then records user feedback on intervention

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 and described above.
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. Results of Pathway Choice
    a. Correct
    b. Minor Deviations/Cycle (describe below)
    c. Major Deviations (describe below)
    d. Comments:
7. Record Errors and Verbalizations:
8. 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(b)(11) Decision Support Intervention

Task 3: Admin User Exports User Feedback

The admin user exports user feedback for evidenced-based DSI. The admin user confirms the file is in computable format with user feedback information

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 and described above.
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. Results of Pathway Choice
    a. Correct
    b. Minor Deviations/Cycle (describe below)
    c. Major Deviations (describe below)
    d. Comments:
7. Record Errors and Verbalizations:
8. 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(b)(11) Decision Support Intervention

Task 4: Admin User Configures User-supplied Predictive DSI and Records / Changes / Access Source Attributes and Then User Triggers User-supplied Predictive DSI

The authorized admin user goes to EHR to select or activate a user-supplied predictive DSI. After user-supplied predictive DSI is selected, the admin user will access source attributes and then record and change them. Then user engages and activates user-supplied predictive DSI and triggers an intervention.

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 and described above.
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. Results of Pathway Choice
    a. Correct
    b. Minor Deviations/Cycle (describe below)
    c. Major Deviations (describe below)
    d. Comments:
7. Record Errors and Verbalizations:
8. 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(b)(11) Decision Support Intervention

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.