EHR Usability Test Report of ASP.MD Inc. Medical Office System (AMOS) version 92. Prepared in accordance with NISTIR 7742 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

ASP.MD Inc Medical Office System (AMOS) v92

Date of Usability Test: 12/26/2018-1/17/2019
Date of Report: 1/20/2019
Report Prepared By: Prakash Pisipati
The New Protocol
Farmingdale, NY
516-321-0312
prakash@newprotocol.com

Table of Contents

EXECUTIVE SUMMARY

A usability test of AMOS v92 was conducted on 12/26/2018-1/17/2019 in Cambridge, MA by Prakash Pisipati. 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 (AMOS V92). The EHR was developed following usability guidelines outlined in

NISTIR 7741 (https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf)

During the usability test, [10] users matching the target demographic criteria served as participants and used AMOS v92 in simulated, but representative tasks.

This study collected performance data on 20 tasks typically conducted on an EHR:

1. Add / register a patient
2. Search for a patient
3. Edit a patient
4. Upload a patients record obtained from an external EHR
5. Import patients medical record document
6. View patient alerts
7.
    a. Add a task to prescribe medication (Tylenol)
    b. Add a task to enter a lab test (CBC)
    c. Add a task to enter a diagnostic test (X-ray)
8. Reconcile a problem in the problem list
9. Add a problem to the patients problem list
10. Reconcile allergies
11. Add an allergy to a medication
12. Reconcile medications
13. Add a medication without prescribing (aspirin)
14. Edit a prescription
15. View medication history
16.
    a. Prescribe a medication
    b. Check drug drug and drug allergy interaction
17. Place an order for a lab test (cbc)
18. Place an order for a diagnostic test
19. Change order status of an existing order
20. Add an implantable device in the device list

During the 20 minute one-on-one usability test, each participant was greeted by the administrator and asked to review and sign an informed consent/release form (included in Appendix 3); they were instructed that they could withdraw at any time. Participants did not have prior experience with the EHR.[4] The administrator

introduced the test, and instructed participants to complete a series of tasks (given one at a time) using AMOS v92. During the testing, the administrator 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.

[4] Basic training comparable to what actual users receive on the same tasks was provided to each participant before the initiation of testing. The users were showed how to perform each task, and then asked to perform each task individually before initiation of testing.

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 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 AMOS v92. Following is a summary of the performance and rating data collected on AMOS v92. Testing was performed by PP.

ID	Task Description	Success Mean(%)	Success-Std Dev(%)	Path Dev-Obs #	Path Dev-Opt #	Time-Mean(s)	Time-Std Dev (s)	Time Dev - Mean Obs Secs	Time Dev-Mean Opt Secs	Errs Mean(%)	Errs-Std Dev (%)	Rating-Scale Type	Rating	Rating Std Dev
1	Add/register a new patient	100	0	7	7	58.7	17.3	13.7	45	0	0	Likert	5.0	0.0
2	Search for a patient	100	0	4	3	10.4	4.8	5.4	5	0	0	Likert	4.9	0.3
3	Edit a patient	100	0	4	3	13.3	7.2	3.3	10	0	0	Likert	4.7	0.7
4	Upload the patient's medical record document obtained from an external EHR	100	0	4	3	22.3	20.1	7.3	15	0	0	Likert	4.8	0.4
5	Import patient's medical record and document	100	0	3	3	22.3	20.1	7.3	15	0	0	Likert	4.9	0.3
6	View patient alerts	100	0	6	5	15.6	9.7	11.6	4	0	0	Likert	4.8	0.6
7	Add a task to precribe medication (Tylenol); Add a task to enter a lab test (CBC); Add a task to enter a diagnostic test (X-ray)	100	0	20	7	32.6	32.4	17.6	15	0	0	Likert	4.4	1.1
8	Reconcile a problem in the problem list	100	0	4	3	13.6	10.6	8.6	5	0	0	Likert	4.7	0.7
9	Add a problem to the patient's problem list (Diabetes)	100	0	5	4	19.2	15.0	14.2	5	0	0	Likert	4.7	0.7
10	Reconcile allergies	100	0	3	3	11.1	4.7	6.1	5	0	0	Likert	5.0	0.0
11	Add an allergy for a medication (Motrin/Abd bleeding)	100	0	6	5	20.2	8.8	5.2	15	0	0	Likert	4.7	0.7

														0.3
12	Reconcile medications	100	0	3	3	15.5	11.1	5.5	10	0	0	Likert	4.9	0.7
13	Add a medication **without prescribing** (Aspirin)	100	0	5	5	36.5	22.7	11.5	25	0	0	Likert	4.7	0.6
14	Edit a prescription	100	0	5	4	25.7	35.4	5.7	20	0	0	Likert	4.8	0.0
15	View medication history	100	0	2	2	4.4	2.7	-5.6	10	0	0	Likert	5.0	0.3
16	Prescribe a medication (Motrin); Check drug-drug & drug-allergy interaction	100	0	8	8	37.9	14.0	7.9	30	0	0	Likert	4.9	0.4
17	Place an order for a lab test (CBC)	100	0	6	6	27.2	14.3	12.2	15	0	0	Likert	4.8	0.6
18	Place an order for diagnostic test (X-ray)	100	0	5	5	18.1	12.8	3.1	15	0	0	Likert	4.8	0.8
19	Change order status of an existing order (from Pending to Completed)	100	0	6	5	25.2	30.7	20.2	5	0	0	Likert	4.6	0.0
20	Add an implantable device in the Device list	100	0	3	3	13.7	4.8	3.7	10	0	0	Likert	5.0	

The results from the System Usability Scale (shown below) scored the subjective satisfaction with the system based on performance with these tasks to be: 84.5.[5]

USABILITY	Questionnaire Info	Rating (1-5)	Adjusted	1 = Strongly Agree	5 = Strongly Disagree
1	I think that I would like to use this system frequently	1.6	1.6		
2	I found the system unnecessarily complex	3.9	2.1		
3	I thought the system was easy to use	1.3	1.3		
4	I think that I would need the support of a technical person to be able to use this system	4.2	1.8		
5	I found the various functions in this system were well integrated	1.5	1.5		
6	I thought there was too much inconsistency in this system	4.7	1.3		
7	I would imagine that most people would learn to use this system very quickly	1.6	1.6		
8	I found the system very cumbersome to use	4.1	1.9		
9	I felt confident using the system	1.4	1.4		
10	I needed to learn a lot of things before I could get going with this system	4.3	1.7		
				Normalized	
			1.62	**84.5**	

In addition to the performance data, the following qualitative observations were made:

- Major findings

  All users agreed that the EHR was easy to navigate and was user friendly. Those who had used other EHR systems previously found AMOS to be more user friendly than those systems.

- Areas for improvement

  Allergies:

  Highlight recently added allergy.

  When required field not entered, highlight it in red (e.g reaction).

  ~~Required fie~~lds should be bold

Systemwide:

Consistent terms; inactivate or deactivate
Use tabs instead of radio buttons for inactivate/deactivate
Show spinner during delays

[5] See Tullis, T. & Albert, W. (2008). Measuring the User Experience. Burlington, MA: Morgan Kaufman (p. 149). Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

INTRODUCTION

The EHRUT tested for this study was ASP.MD Medical Office System (AMOS) v92. Designed to present medical information to healthcare providers in multispecialty outpatient office settings AMOS v92 consists of totally web based practice management and electronic medical record functionality. 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 (AMOS V92). To this end, measures of effectiveness, efficiency and user satisfaction, such as time on task, deviation from suggested path, and errors, were captured during the usability testing.

METHOD

PARTICIPANTS

A total of 10 participants were tested on AMOS v92. Participants in the test were doctors, PAs, students, physical therapy students, and medical office managers. Participants were recruited by Prakash Pisipati and were compensated $200 for their time. In addition, participants had no direct connection to the development of or organization producing AMOS V92. Participants were not from the testing or supplier organization. Participants were given the opportunity to have the same orientation and level of training as the actual end users would have received.

For the test purposes, end-user characteristics were identified and translated into a recruitment screener used to solicit potential participants; an example of a screener is provided in Appendix 1.

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.

PARTICIPANT DEMOGRAPHICS

Following is a high-level overview of the participants in this study.

Use	Date	ID	Sex	Age	Ed	Role	Prof Exp	Cmp Exp	Prd Exp	Needs
Internal Medicine	12/26/2018	1-MP	F	20-29	High school graduate	College Student	1	120	0	None
Internal Medicine	12/27/2018	2-KK	F	50-59	Bachelor's Degree	Office Manager	25	60	0	None
Internal Medicine	12/27/2018	3-CC	F	40-49	Bachelor's Degree	RN	25	120	0	None
Internal Medicine	1/4/2019	4-RO	F	20-29	Master's Degree	PA	3	120	0	None
Internal Medicine	1/14/2019	5-MH	M	20-29	Master's Degree	PA	4	120	0	None
Internal Medicine	15-Jan	6-JO	F	30-39	Doctorate degree	MD	2	120	0	None
Internal Medicine	1/15/2019	7-CP	F	30-39	Bachelor's Degree	Office Manager	5	120	0	None
Internal Medicine	1/16/2019	8-MK	F	30-39	Doctorate degree	MD	1	120	0	None
Internal Medicine	1/16/2019	9-JK	M	30-39	Master's Degree	PA	5	120	0	None
Internal Medicine	1/17/2019	10-LR	F	40-49	Bachelor's Degree	RN	22	120	0	None

10 participants (matching the demographics in the section on Participants) were recruited and 10 participated in the usability test. 0 participants failed to show for the study.

Participants were scheduled for 2 hour sessions with

1 hour in between each session for debrief by the administrator, and

to reset systems to proper test conditions. A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics as provided by the recruiting firm.

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 1 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 Section 3.9 on Usability Metrics.

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, including:

1. Add/register a new patient
2. Search for a patient
3. Edit a patient
4. Upload the patient's medical record document obtained from an external EHR
5. Import patient's medical record and document
6. View patient alerts
7. Add a task to precribe medication (Tylenol); Add a task to enter a lab test (CBC); Add a task to enter a diagnostic test (X-ray)
8. Reconcile a problem in the problem list
9. Add a problem to the patient's problem list (Diabetes)
10. Reconcile allergies
11. Add an allergy for a medication (Motrin/Abd bleeding)
12. Reconcile medications
13. Add a medication **without prescribing** (Aspirin)
14. Edit a prescription
15. View medication history
16. Prescribe a medication (Motrin); Check drug-drug & drug-allergy interaction
17. Place an order for a lab test (CBC)
18. Place an order for diagnostic test (X-ray)
19. Change order status of an existing order (from Pending to Completed)
20. Add an implantable device in the Device list

Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users.[6] Tasks

should always be constructed in light of the study objectives.

PROCEDURES

Upon arrival, participants were greeted; their identity was verified and matched with a name on the participant schedule. Participants were then assigned a participant ID.[7]

Each participant reviewed and signed an informed consent and release form (See Appendix 3). A representative from the test team witnessed the participant's signature.

The tasks were administered by Prakash Pisipati.

The administrator moderated the session including administering instructions and tasks. The administrator also monitored task times, obtained post-task rating data, and took notes on participant comments. The administrator also served as the data logger and took notes on task success, path deviations, number and type of errors, and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below in Section 3.9.

Following the session, the administrator gave the participant the post-test questionnaire (e.g., the System Usability Scale, see Appendix 5), compensated them for their time, and thanked each individual for their participation.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

Participants were thanked for their time and compensated. Participants signed a receipt and acknowledgement form (See Appendix 6) indicating that they had received the compensation.

TEST LOCATION

The test facility included a waiting area and a quiet testing room with a table, computer for the participant, and recording computer for the administrator. Only the participant and administrator were in the test room. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, and visible to the participants.

TEST ENVIRONMENT

AMOS v92 would be typically be used in a healthcare office or facility. In this instance, the testing was conducted in an office. For testing, the computer used a PC running Windows. The participants used a mouse and keyboard when interacting with AMOS v92.

AMOS v92 used a 24” display at maximum resolution and color settings.

The application was set up by the vendor according to the vendor’s documentation describing the

system set-up and preparation. The application itself was running on a Microsoft platform using a test database] on a LAN connection. Technically, the system performance (i.e., response time) was representative to what actual users would experience in a field implementation. Additionally, participants were instructed not to change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

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

1. Informed Consent
2. Moderator's Guide
3. Post-test Questionnaire
4. Incentive Receipt and Acknowledgment Form

Examples of these documents can be found in Appendices 3-6 respectively. The Moderator's Guide was devised so as to be able to capture required data.

The participant's interaction with AMOS v92 was captured and recorded digitally with screen capture software running on the test machine. A web camera recorded each participant's facial expressions synced with the screen capture, and verbal comments were recorded with a microphone. [8] The test session were electronically transmitted to a nearby observation room where the data logger observed the test session.

The administrator reads the following instructions aloud to the each participant (also see the full moderator's guide in Appendix [B4]):

> *Thank you for participating in this study. Your input is very important. Our session today will last about 10 minutes. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely. Please note that we are not testing you we are testing the system, therefore if you have difficulty all this means is 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.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given time 15 minutes) to explore the system and make comments. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you* not *talk aloud or verbalize while you are doing the tasks.*[9] *I will ask you your impressions about the task once*

you are done.

Participants were then given 20 tasks to complete. Tasks are listed in the moderator's guide in Appendix [B4].

[8] There are a variety of tools that record screens and transmit those recordings across a local area network for remote observations.

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 AMOS v92 by measuring participant success rates and errors
2. Efficiency of AMOS v92 by measuring the average task time and path deviations

[9] Participants should not use a think-aloud protocol during the testing. Excessive verbalization or attempts to converse with the moderator during task performance should be strongly discouraged. Participants will naturally provide commentary, but they should do so, ideally, after the testing. Some verbal commentary may be acceptable between tasks, but again should be minimized by the moderator.

3. Satisfaction with AMOS v92 by measuring ease of use ratings

DATA SCORING

The following table (Table [x]) details how tasks were scored, errors evaluated, and the time data analyzed.[10]

Measures	Rationale 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 30 seconds then allotted task time performance was 30 * 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.[11] 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: Task Deviations	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. 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.

[10] An excellent resource is Tullis, T. & Albert, W. (2008). Measuring the User Experience. Burlington, MA: Morgan Kaufman. Also see www.measuringusability.com

[11] Errors have to be operationally defined by the test team prior to testing.

	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. [12] 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 AMOS v92 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 5. [13]

Table [x]. Details of how observed data were scored.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above.

[12] See Tedesco and Tullis (2006) for a comparison of post-task ratings for usability tests. Tedesco, D. & Tullis, T. (2006) A comparison of methods for eliciting post-task subjective ratings in usability testing. *Usability Professionals association Conference*, June 12 – 16, Broomfield, CO.

[13] The SUS survey yields a single number that represents a composite measure of the overall perceived usability of the system. SUS scores have a range of 0 to 100 and the score is a relative benchmark that is used against other iterations of the system.

The usability testing results for AMOS v92 are detailed in the table below.

USABILITY	Questionnaire Info	Rating (1-5)	Adjusted
1	I think that I would like to use this system frequently	1.6	1.6
2	I found the system unnecessarily complex	3.9	2.1
3	I thought the system was easy to use	1.3	1.3
4	I think that I would need the support of a technical person to be able to use this system	4.2	1.8
5	I found the various functions in this system were well integrated	1.5	1.5
6	I thought there was too much inconsistency in this system	4.7	1.3
7	I would imagine that most people would learn to use this system very quickly	1.6	1.6
8	I found the system very cumbersome to use	4.1	1.9
9	I felt confident using the system	1.4	1.4
10	I needed to learn a lot of things before I could get going with this system	4.3	1.7
			1.62

Adjusted to 1-100 Scale **84.5**

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be: 84.5. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.[15]

---

[15] See Tullis, T. & Albert, W. (2008). Measuring the User Experience. Burlington, MA: Morgan Kaufman (p. 149).

EFFECTIVENESS

AMOS v92 ranked above average (82 in SUS across the board, 4.5/5 in specific tasks) for usability and ease of tasks respectively.

EFFICIENCY

AMOS v92 varied with respect to efficiency of tasks. Most tasks were completed within “reasonable” times, however, 2 tasks required longer than expected; dismissing an alert, and ordering a lab. With regard to dismissing an alert, we think that hesitation in dismissal accounted for the delay, and that with repeated use, users would become familiar with the concept and dismissals would be more rapid. With regard to ordering labs, the order entry interface is one of the more complex interfaces. There might be a few measures which could be taken to simplify this interface to speed order entry. Further evaluation might be necessary, however, to determine if other orders outside of the order evaluated are also delayed.

SATISFACTION

Overall satisfaction ranking was 82, considered above average. Lower scores were reported with the need for assistance and comfort with the system; however, we expect any first time users to experience a need for assistance and some level of discomfort with the new system based on uncertainty.

MAJOR FINDINGS

Participants ranked the system 4.5/5 for ease of use on all tasks overall. Participants did not universally have difficulty with any specific tasks. Some tasks did take longer than would be desired, including dismissing alerts and ordering labs. Some tasks also considered relatively complex, such as prescribing medications, did seem to be managed by all participants without significant difficulty.

On verbal followup with participants, some felt that there was some lack of consistency within the interface; for example that inactivation in one place might be represented by an X, but by an I in another. They also indicated the desire for system activity indicators while waiting for some actions to occur.

AREAS FOR IMPROVEMENT

While all tasks were competed in a reasonable time, there is always room to shave even more time off task completion. We also identified a number of areas where users felt uncomfortable or confused (alerts and orders). Finally, we agree that there are some inconsistencies across the system. We know from experience that users quickly adapt to these quirks, however, it is our goal to eliminate them nonetheless to make use easier for first time users.

APPENDICES

1: Recruiting screener

2: Participant demographics

3: Non-Disclosure Agreement (NDA) and Informed Consent Form

4: Moderator's Guide

5: System Usability Scale Questionnaire

6: Incentive receipt and acknowledgment form

Appendix 1: RECRUITING SCREENER

This is a sample of the recruiting screener used for this study:

Recruiting Script for Recruiting Firm

Hello, my name is _____, calling from *[Insert name of recruiting firm]*. We are recruiting individuals to participate in a usability study for an electronic health record. We would like to ask you a few questions to see if you qualify and if would like to participate. This should only take a few minutes of your time. This is strictly for research purposes. If you are interested and qualify for the study, you will be paid to participate. Can I ask you a few questions?

Customize this by dropping or adding questions so that it reflects your EHR's primary audience

1. [If not obvious] Are you male or female? [Recruit a mix of participants]
2. Have you participated in a focus group or usability test in the past *xx* months? [If yes, Terminate]
3. Do you, or does anyone in your home, work in marketing research, usability research, web design […etc.]? [If yes, Terminate]
4. Do you, or does anyone in your home, have a commercial or research interest in an electronic health record software or consulting company? [If yes, Terminate]
5. Which of the following best describes your age? [23 to 39; 40 to 59; 60 - to 74; 75 and older] [Recruit Mix]
6. Which of the following best describes your race or ethnic group? [e.g., Caucasian, Asian, Black/African-American, Latino/a or Hispanic, etc.]
7. Do you require any assistive technologies to use a computer? [if so, please describe]

Professional Demographics *Customize this to reflect your EHR's primary audience*

8. What is your current position and title? (Must be healthcare provider)
   - ☐ RN: Specialty _____
   - ☐ Physician: Specialty _____
   - ☐ Resident: Specialty _____
   - ☐ Administrative Staff
   - ☐ Other [Terminate]

9. How long have you held this position?
10. Describe your work location (or affiliation) and environment? (Recruit according to the intended users of the application) [e.g., private practice, health system, government clinic, etc.]
11. Which of the following describes your highest level of education? [e.g., high school graduate/GED, some college, college graduate (RN, BSN), postgraduate (MD/PhD), other (explain)]

Computer Expertise *Customize this to reflect what you know about your EHR's audience*

12. Besides reading email, what professional activities do you do on the computer? [e.g., access EHR, research; reading news; shopping/banking; digital pictures; programming/word processing, etc.] [If no computer use at all, Terminate]
13. About how many hours per week do you spend on the computer? [Recruit according to the demographics of the intended users, e.g., 0 to 10, 11 to 25, 26+ hours per week]
14. What computer platform do you usually use? [e.g., Mac, Windows, etc.]
15. What Internet browser(s) do you usually use? [e.g., Firefox, IE, AOL, etc.]
16. In the last month, how often have you used an electronic health record?
17. How many years have you used an electronic health record?
18. How many EHRs do you use or are you familiar with?
19. How does your work environment patient records? [Recruit according to the demographics of the intended users]
    - ☐ On paper
    - ☐ Some paper, some electronic
    - ☐ All electronic

Contact Information *If the person matches your qualifications, ask*

Those are all the questions I have for you. Your background matches the people we're looking for. For your participation, you will be paid $200.

May I get your contact information?

- Name of participant:
- Address:
- City, State, Zip:
- Daytime phone number:
- Evening phone number:
- Alternate [cell] phone number:
- Email address:

Before your session starts, we will ask you to sign a release form allowing us to videotape your session. The videotape will only be used internally for further study if needed. Will you consent to be videotaped?

This study will take place at 229 3rd St Cambridge, MA 02141. I will confirm your appointment a couple of days before your session and provide you with directions to our office. What time is the best time to reach you?

Appendix 2: PARTICIPANT DEMOGRAPHICS

Following is a high-level overview of the participants in this study.

Gender	**Occupation/Role**	**Years of Experience (Average)**	**Facility Use of EHR**
Men ---------- 3 Women---------- 7 Total (participants)---10	Students---------- 1 RN---------- 3 PA---------- 3 Off. Mgr.---------- 2 MD/DO---------- 1 Total (participants) --10	Students---------- 1 RN---------- 12 PA---------- 4 Off. Mgr.---------- 15 MD/DO ---------- 2	All paper ---------- 0 Some paper, some electronic----------10 All electronic---------- 0 Total(participants)------10

As an appendix to the report, the full participant breakdown (de-identified) should be included.

Appendix 3: NON-DISCLOSURE AGREEMENT AND INFORMED CONSENT FORM

Sample of the Non-Disclosure Used by Us

Non-Disclosure Agreement

THIS AGREEMENT is entered into as of _____, *2018*, between _____ ("the Participant") and the testing organization *Test Company* located at *Address*.

The Participant acknowledges his or her voluntary participation in today's usability study may bring the Participant into possession of Confidential Information. The term "Confidential Information" means all technical and commercial information of a proprietary or confidential nature which is disclosed by *Test Company*, or otherwise acquired by the Participant, in the course of today's study.

By way of illustration, but not limitation, Confidential Information includes trade secrets, processes, formulae, data, know-how, products, designs, drawings, computer aided design files and other computer files, computer software, ideas, improvements, inventions, training methods and materials, marketing techniques, plans, strategies, budgets, financial information, or forecasts.

Any information the Participant acquires relating to this product during this study is confidential and proprietary to *Test Company* and is being disclosed solely for the purposes of the Participant's participation in today's usability study. By signing this form the Participant acknowledges that s/he will receive monetary compensation for feedback and will not disclose this confidential information obtained today to anyone else or any other organizations.

Participant's printed name: _____

Signature: _____ **Date:** _____

Informed Consent

Test Company 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. At the conclusion of the test, you will be compensated for your time.

Agreement

I understand and agree that as a voluntary participant in the present study conducted by *Test Company* I am free to withdraw consent or discontinue participation at any time. I understand and agree to participate in the study conducted and videotaped by the *Test Company*.

I understand and consent to the use and release of the videotape by *Test Company*. I understand that the information and videotape 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 videotape and understand the videotape may be copied and used by *Test Company* 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 with outside of *Test Company* and *Test Company's* client. 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 4: EXAMPLE MODERATOR'S GUIDE

The template below was followed by our administrator.

AMOS V92 Usability Test
Moderator's Guide

Administrator _____

Data Logger _____

Date _____ **Time** _____

Participant # _____

Location _____

Prior to testing
- Confirm schedule with Participants
- Ensure AMOS V92 lab environment is running properly
- Ensure lab and data recording equipment is running properly

Prior to each participant:
- Reset application
- Start session recordings with *tool*

Prior to each task:
- Reset application to starting point for next task

After each participant:
- End session recordings with *tool*

After all testing
- Back up all video and data files

Orientation (*X* minutes)

Thank you for participating in this study. Our session today will last about 20 minutes. During that time you will take a look at an electronic health record system.

I will ask you to complete a few tasks using this system and 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.

I did not have any involvement in its creation, so please be honest with your opinions.

The product you will be using today is the ASP.MD Medical Office System v92. Some of the data may not make sense as it is placeholder data.

We are recording the audio and screenshots of our session today. 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?

Preliminary Questions (*X* minutes)

What is your job title / appointment?

How long have you been working in this role?

What are some of your main responsibilities?

Tell me about your experience with electronic health records.

The following template was followed for each task in the evaluation:

Take the participant to the starting point for the task.

Task is described to the subject here.

Success:

☐ Easily completed
☐ Completed with difficulty or help :: Describe below
☐ Not completed

Comments:

Task Time: _____ Seconds

Optimal Path: *Screen A → Screen B → Drop Down B[1] → "OK" Button → Screen X…*

☐ Correct
☐ Minor Deviations / Cycles :: Describe below
☐ Major Deviations :: Describe below

Comments:

Observed Errors and Verbalizations:

Comments:

Rating:

Overall, this task was: _

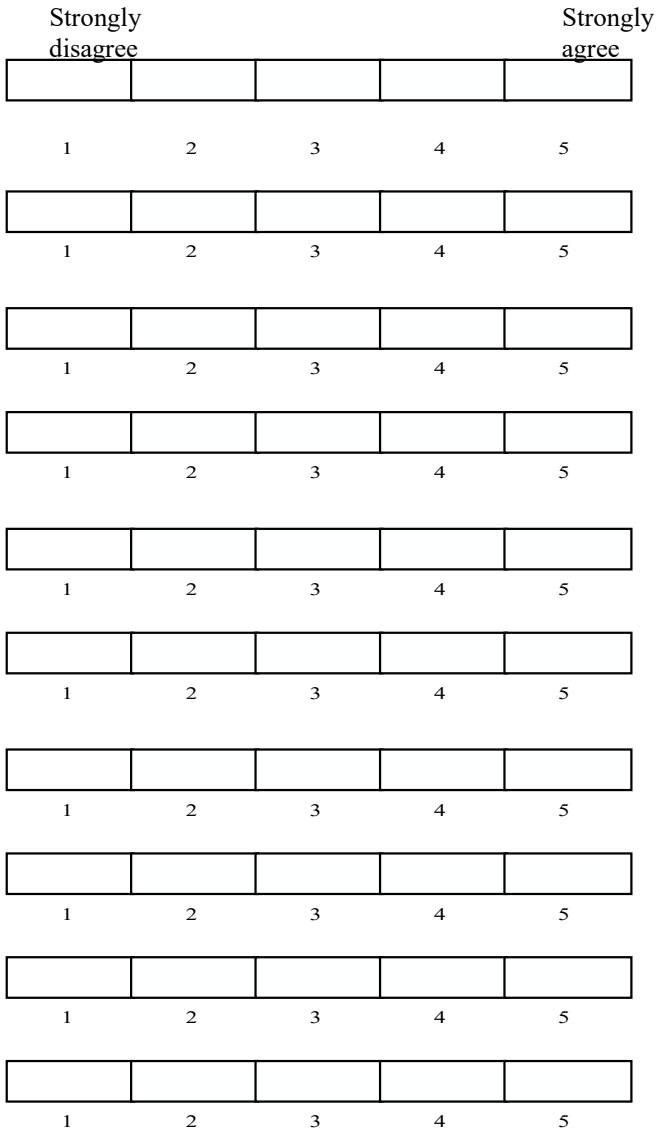
Show participant written scale: "Very Easy" (1) to "Very Difficult" (5)

Administrator / Notetaker Comments:

The template above was applied to all tasks in the evalution. Details are attached in test evaluation document.

Appendix 5: SYSTEM USABILITY SCALE QUESTIONNAIRE

*In 1996, Brooke published a "*low-cost usability scale that can be used for global assessments of systemsusability*" known as the System Usability Scale or SUS.*[16]
Lewis and Sauro (2009) and others have elaborated on the SUS over the years. Computation of the SUS score can be found in Brooke's paper, in at
http://www.usabilitynet.org/trump/documents/Suschapt.doc *or in Tullis and Albert (2008).*



[16] Brooke, J.: SUS: A "quick and dirty" usability scale. In: Jordan, P. W., Thomas, B., Weerdmeester, B. A., McClelland (eds.) *Usability Evaluation in Industry* pp. 189--194. Taylor & Francis, London, UK (1996). SUS is copyrighted to Digital Equipment Corporation, 1986.

Lewis, J R & Sauro, J. (2009) "The Factor Structure Of The System Usability Scale." in *Proceedings of the Human Computer Interaction International Conference (HCII 2009), San Diego CA, USA*

Acknowledgement of Receipt

I hereby acknowledge receipt of *$_____* for my participation in a research study run by *Test Company.*

Printed Name: _____

Address: _____

_____

Signature: _____ Date: _____

Usability Researcher: _____

Signature of Usability Researcher: _____

Date: _____

Witness: _____

Witness Signature: _____

Date: _____

NIST 7742 EHR Usability Test Report of ASP.MD Medical Office System Version 92

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

ASP.MD Medical Office System (AMOS) v92

Date of Usability Test: 2025/02/03
Date of Report: 2025/02/05
Report Prepared By: Akhil Nair , Research Head, Yellow slice
+91-9175279654
akhil.n@yellowslice.in
Maple woodz, Wagholi, Pune, India

Table of Contents

EXECUTIVE SUMMARY

A usability test of ASP.MD Medical Office System version 92 was conducted on 2025/02/03 in by Yellowslice. 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).

The EHR was developed following usability guidelines outlined in NISTIR 7741, NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records (https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf)

During the usability test, 10 subjects matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

This study collected performance data on [7] tasks typically

conducted on an EHR:

1. Looking up the patient and opening alerts
2. Checking Health report
3. Overriding and alert
4. Checking the supporting documents on alerts
5. Adding feedback to alerts
6. Checking the source information of the alerts
7. Running alert feedback report

During the 15-minute one-on-one usability test, each participant was greeted by the administrator and asked to review and sign an informed consent/release form (included in Appendix 3); they were instructed that they could withdraw at any time. Participants did not have any prior experience with the EHR and thus were given basic training on the platform to give them an understanding of the activities and functionalities. The administrator introduced the test and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test and, along with the data logger(s) recorded user performance data electronically. The administrator did not give the participant much

assistance in how to complete the task.

Participant screens 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

Task Identifier	Task Description	Task Success - Mean (%)	Task Success - Std Dev (%)	Task Path Deviation - Observed #	Task Path Deviation - Optimal #	Task Time - Mean (seconds)	Task Time - Standard Deviation (seconds)	Task Time Deviation - Observed Seconds	Task Time Deviation - Optimal Seconds	Task Errors Mean (%)	Task Errors - Std Dev (%)	Task Rating - Scale Type	Task Rating	Task Rating - Standard Deviation	Participant Identifiers	UCD Process Selected	UCD Process Details
a.1	Looking up the patient and opening alerts - Tests criterion 170.315(b)(11)	100	0	5	4	18	9	11	8	0	0	Likert	4.25	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741	was applied
a.2	Checking Health report - Tests criterion 170.315(b)(11)	100	0	1	1	7	3	3	4	0	0	Likert	4.25	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741	was applied
a.3	Overriding an alert - Tests criterion 170.315(b)(11)	100	0	6	5	14	9	7	6	0	0	Likert	4.00	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741	was applied
a.4	Checking the supporting documents on alert - Tests criterion 170.315(b)(11)	100	0	1	1	4	2	1	3	0	0	Likert	4.75	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741	was applied
a.5	Adding feedback to alert - Tests criterion 170.315(b)(11)	100	0	2	2	9	5	4	5	0	0	Likert	4.75	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741	was applied
a.6	Checking the Source Information of alert - Tests criterion 170.315(b)(11)	100	0	1	1	6	2	2	4	0	0	Likert	4.37	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741	was applied

a.7	Running alert feedback report - Tests criterion 170.315(b)(11)	60	49	5	4	15	7	6	9	0	0	Likert	4.15	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741	NISTIR 7741 was applied

- Time to complete the tasks
- Number and types of errors
- Path deviations
- 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 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.

The results from the Likert Scale scored the subjective satisfaction with the system based on performance with these tasks to be: **4.36**

In addition to the performance data, the following qualitative observations were made:

- Major findings
    - Participants did not universally have difficulty with any specific tasks. Some tasks did take longer than desired, including searching up the Patient opening alerts and running reports for alert feedback.
    - On verbal follow-up with participants, some felt that the system looked outdated and identifying the visual elements is not easy as it is a mental task to navigate. Participants felt confident to complete the tasks once again, however, each one felt that without a quick guide, it would be very complex to navigate and complete the journey.
- Areas for improvement
    - Participants agreed that they would need training on the system prior to using it, indicating the system is slightly complex to use on their own.

Some basic changes like making the buttons (A, O, S) more descriptive (Alerts, Overriding, Source Information) will give more clarity to the users and they will be able to explore things on their own.

INTRODUCTION

The EHRUT tested for this study was ASP.MD Medical Office System (AMOS) version 92. Designed to present medical information to healthcare providers in multispecialty outpatient office settings the EHRUT consists of totally web-based practice management and electronic medical record functionality. 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 on task, deviation from suggested path, and errors, were captured during the usability testing.

METHOD

PARTICIPANTS

A total of 10 participants were tested on the EHRUT(s). Participants in the tests were Doctors, Medical Students, and Healthcare Consultants along with 1 product manager. Participants were recruited by Yellowslice and were compensated for their time. In addition, participants had no direct connection to the development of or organization producing the EHRUT(s). Participants were not from the testing or supplier organization. Participants were given the opportunity to have the same orientation and level of training as the actual end users would have received.

For the test purposes, end-user characteristics were identified and

translated into a recruitment screener used to solicit potential

participants; an example of a screener is provided in Appendix [1].

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.

Participant Identifier	Participant Gender	Participant Age	Participant Education	Participant Occupation/Role	Participant Professional Experience (Months)	Participant Computer Experience (Months)	Participant Product Experience	Participant Assistive Technology Needs
P1	Female	20-29	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Doctor	60.00	56.00	0.00	No
P2	Female	20-29	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Doctor	12.00	60.00	0.00	No
P3	Male	40-49	Bachelor's degree	Homeopathic Consultant	240.00	180.00	0.00	No
P4	Male	40-49	Master's degree	Health Claims Team	156.00	108.00	0.00	No
P5	Male	20-29	Master's degree	Assistant Professor	24.00	240.00	0.00	No
P6	Male	40-49	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Doctor	288.00	252.00	0.00	No
P7	Female	40-49	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Doctor	240.00	120.00	0.00	No
P8	Male	30-39	Bachelor's degree	Product Manager	192.00	180.00	0.00	No
P9	Female	20-29	Doctorate degree (e.g., MD, DNP,	Doctor	24.00	60.00	0.00	No

			DMD, PhD)					
P10	Male	20-29	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Doctor	36.00	120.00	0.00	No

10 participants (matching the demographics in the section on Participants) were recruited and 10 participated in the usability test. 0 participants failed to show up for the study.

Participants were scheduled for 30-minute sessions. A spreadsheet was used to keep track of the participant's schedule and included each participant's demographic characteristics as provided by the recruiting firm.

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 1 EHR. Each participant used the system from a different 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 satisfaction ratings of the system

Additional information about the various measures can be found in Section 3.9 on Usability Metrics

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, including:

1. Looking up the patient and opening alerts
2. Checking Health report
3. Overriding and alert
4. Checking the supporting documents on alerts
5. Adding feedback to alerts
6. Checking the source information of the alerts
7. Running alert feedback report

Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Tasks should always be constructed in light of the study objectives.

PROCEDURES

Upon joining the call, participants were greeted; their identity was verified and matched with a name on the participant schedule. Participants were then assigned a participant ID. Each participant reviewed and signed an informed consent and release form (See Appendix 3).

To ensure that the test ran smoothly, two staff members participated in this test, the usability administrator and the data logger. The administrator moderated the session including administering instructions and tasks. The administrator also monitored task times, obtained post-task rating data, and took notes on participant comments. A second person served as the data logger and took notes

on task success, path deviations, number and type of errors, and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think-aloud technique

The participants were given a basic walkthrough of the platform in the beginning of the session. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below in Section 3.9.

Following each of the tasks, the users were asked to rate the difficulty of each of the tasks, once the session was completed, the participants were compensated for their time, and thanked each individual for their participation.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, were recorded into a spreadsheet.

Participants were thanked for their time and compensated. Participants signed a receipt and acknowledgment form (See Appendix 6) indicating that they had received the compensation.

TEST LOCATION

The tests were performed over zoom with the participants doing the testing in a location of their choice. These locations were either the participants homes or their offices.

TEST ENVIRONMENT

The EHRUT would typically be used in a healthcare office or facility or might be accessed from a remote location such as remote office or home.

In this instance, the testing was conducted online over a Zoom Call. For testing, the computer used a Macbook to conduct the test with MacOS.

The participants used a laptop when interacting with the EHRUT.

Technically, the system performance (i.e., response time) was representative of what actual users would experience in a field implementation. Additionally, participants were instructed not to change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

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

1. Informed Consent
2. Moderator's Guide

3. Incentive Receipt and Acknowledgment Form

Examples of these documents can be found in Appendices 3-6 respectively. The Moderator's Guide was devised so as to be able to capture the required data.

The participant's interaction with the EHRUT was captured and recorded digitally with screen capture software running on the test machine. The data logger on the same call records the data.

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to each participant (also see the full moderator's guide in Appendix [B4]):

> *Thank you for participating in this study. Your input is very important. Our session today will last about 10 minutes. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely. Please note that we are not testing you we are testing the system, therefore if you have difficulty all this means is 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.

Following the procedural instructions, participants were shown the EHR and as their first task, were given time few minutes) to explore the system and make comments. Once this task was complete, the administrator gave the following instructions:

For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not *talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given 7 tasks to complete. Tasks are listed in the moderator's guide in Appendix [B4].

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 [EHRUT] by measuring participant success rates and errors
2. Efficiency of [EHRUT] by measuring the average task time and path deviations

3. Satisfaction with [EHRUT] by measuring ease of use ratings

DATA SCORING

The following table (Table 1) details how tasks were scored, errors evaluated, and the time data analyzed.[10]

Measures	Rationale 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 [x] seconds then allotted task time performance was [x * 1.25] seconds. 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.[11] 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: Task Deviations	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. 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. A common convention is that average ratings for systems judged easy to use should be 3.3 or above.

Table 1. Details of how observed data were scored.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above.

The usability testing results for the EHRUT are detailed below (see Table [x])[14]. The results should be seen in light of the objectives and goals outlined in Section 3.2 Study Design. The data should yield actionable results that, if corrected, yield material, positive impact on user performance.

Task Identifier	Task Description	Task Success - Mean (%)	Task Success - Std Dev (%)	Task Path Deviation - Observed #	Task Path Deviation - Optimal #	Task Time - Mean (seconds)	Task Time - Standard Deviation (seconds)	Task Time Deviation - Observed Seconds	Task Time Deviation - Optimal Seconds	Task Errors Mean (%)	Task Errors - Std Dev (%)	Task Rating - Scale Type	Task Rating	Task Rating - Standard Deviation	Participant Identifiers	UCD Process Selected	UCD Process Details
a.1	Looking up the patient and opening alerts - Tests criterion 170.315(b)(11)	100	0	5	4	18	9	11	8	0	0	Likert	4.25	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741 was applied	
a.2	Checking Health report - Tests criterion 170.315(b)(11)	100	0	1	1	7	3	3	4	0	0	Likert	4.25	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741 was applied	
a.3	Overriding an alert - Tests criterion 170.315(b)(11)	100	0	6	5	14	9	7	6	0	0	Likert	4.00	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741 was applied	
a.4	Checking the supporting documents on alert - Tests criterion 170.315(b)(11)	100	0	1	1	4	2	1	3	0	0	Likert	4.75	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741 was applied	
a.5	Adding feedback to alert - Tests criterion 170.315(b)(11)	100	0	2	2	9	5	4	5	0	0	Likert	4.75	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741 was applied	
a.6	Checking the Source Information of alert - Tests criterion 170.315(b)(11)	100	0	1	1	6	2	2	4	0	0	Likert	4.37	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741 was applied	

a.7	Running alert feedback report - Tests criterion 170.315(b)(11)	60	49	5	4	15	7	6	9	0	0	Likert	4.15	5.00	P1;P2;P3;P4;P5;P6;P7;P8;P9;P10	NISTIR 7741	NISTIR 7741 was applied

The results from the task rating suggested that users found the tasks easy:

4.36 Rating. A common convention is that average ratings for systems

judged easy to use should be 3.3 or above.

DISCUSSION OF THE FINDING

EFFECTIVENESS

The average rating for all tasks combined is 4.36/5. This indicates that most of the users felt the tasks were easy to perform.

EFFICIENCY

The EHRUT varied with respect to the efficiency of tasks. Most tasks were completed close to the optimal time required, however, 3 tasks took a few extra seconds compared to the optimal time. The labeling of the buttons (A, O, S, etc) is not very clear to the new user, as it doesn't give them any clarity on what that button does. Without training people won't be able to use the platform as intended.

SATISFACTION

Participants ranked the system 4.36/5 for ease of use on all tasks overall. This indicates that most of the users felt the tasks were easy to perform, however, we expect any first-time users to experience a need for assistance and some level of discomfort with the new system based on uncertainty.

MAJOR FINDINGS

Participants did not universally have difficulty with any specific tasks. Some tasks did take longer than desired, including looking up the patient opening alerts and running reports for alert feedback.

On verbal follow-up with participants, some felt that the system looked outdated and identifying the visual elements is not easy as it is a mental task to navigate. Participants felt confident to complete the tasks once again, however, each one felt that without a quick guide, it would be very complex to navigate and complete the journey.

AREAS FOR IMPROVEMENT

Participants agreed that they would need training on the system prior to using it, indicating the system is slightly complex to use on their own. Some basic changes like making the buttons (A, O, S) more descriptive (Alerts, Overriding, Source Information) will give more clarity to the users and they will be able to explore things on their own.

APPENDICES

The following appendices include supplemental data for this usability test report. Following is a list of the appendices provided:

1: Recruiting screener

2: Participant demographics

3: Non-Disclosure Agreement (NDA) and Informed Consent Form

4: Moderator's Guide

5: System Usability Scale Questionnaire

6: Incentive receipt and acknowledgment form

Appendix 1: RECRUITING SCREENER

Following is the screener used to recruit participants

Recruiting Script for Recruiting Firm

Hello, my name is [redacted], calling from YellowSlice. We are recruiting individuals to participate in a usability study for an electronic health record. We would like to ask you a few questions to see if you qualify and if would like to participate. This should only take a few minutes of your time. This is strictly for research purposes. If you are interested and qualify for the study, you will be paid to participate. Can I ask you a few questions?

1. Have you participated in a focus group or usability test in the past *xx* months?
2. Do you, or does anyone in your home, work in marketing research, usability research, web design?
3. Which of the following best describes your age? [23 to 39; 40 to 59; 60 - to 74; 75 and older]
4. Do you require any assistive technologies to use a computer?

Professional Demographics

5. What is your current position and title?
   - ☐ Physician
   - ☐ Other healthcare provider _____
   - ☐ Other healthcare worker_____
   - ☐ Healthcare instructor
   - ☐ Healthcare administrator

6. How long have you held this position?
7. Which of the following describes your highest level of education? [e.g., high school graduate/GED, some college, college graduate (RN, BSN), postgraduate (MD/PhD), other (explain)]

Computer Expertise

8. Besides reading email, what professional activities do you do on the computer? [e.g., access EHR, research; reading news; shopping/banking; digital pictures; programming/word processing, etc.]
9. How does your work environment maintain patient records?
   - ☐ On paper
   - ☐ Some paper, some electronic
   - ☐ All electronic

Contact Information

Those are all the questions I have for you. Your background matches the people we're looking for. [If you are paying participants or offering some form of compensation, mention] For your participation, you will be paid $100.

May I get your contact information?

- Name of participant:
- Contact number:
- Email address:

Before your session starts, we will ask you to sign a release form allowing us to record your session. The recordig will only be used internally for further study if needed. Will you consent to be recorded?

This study will take place over Zoom. I will confirm your appointment a couple of days before your session and provide you with directions to connect to the Zoom session.

Appendix 2: PARTICIPANT DEMOGRAPHICS

Following is a high-level overview of the participants in this study.

Gender	
Men	[6]
Women	[4]
Total (participants)	[10]

Occupation/Role	
RN/BSN	[1]
Physician	[6]
Admin Staff	[3]
Total (participants)	[10]

Years of Experience	
Years experience	[106]
Facility Use of EHR	
All paper	[106]
Some paper, some electronic	[0]
All electronic	[0]
Total (participants)	[106]

As an appendix to the report, the full participant breakdown (de-identified) should be included.

Appendix 3: NON-DISCLOSURE AGREEMENT AND INFORMED CONSENT FORM USED IN THE STUDY

Non-Disclosure Agreement

THIS AGREEMENT is entered into as of _____, *2025*, between _____("the Participant") and the testing organization *YellowSlice* located at Maple woodz, Wagholi, Pune, India.

The Participant acknowledges his or her voluntary participation in today's usability study may bring the Participant into possession of Confidential Information. The term "Confidential Information" means all technical and commercial information of a proprietary or confidential nature which is disclosed by *ASP.MD Inc*, or otherwise acquired by the Participant, in the course of today's study.

By way of illustration, but not limitation, Confidential Information includes trade secrets, processes, formulae, data, know-how, products, designs, drawings, computer aided design files and other computer files, computer software, ideas, improvements, inventions, training methods and materials, marketing techniques, plans, strategies, budgets, financial information, or forecasts.

Any information the Participant acquires relating to this product during this study is confidential and proprietary to *ASP.MD Inc* and is being disclosed solely for the purposes of the Participant's participation in today's usability study. By signing this form the Participant acknowledges that s/he will receive monetary compensation for feedback and will not disclose this confidential information obtained today to anyone else or any other organizations.

Participant's printed name: _____

Signature: _____ **Date:** _____

Informed Consent form used in the study

ASP.MD Inc 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. At the conclusion of the test, you will be compensated for your time.

Agreement

I understand and agree that as a voluntary participant in the present study conducted by *ASP.MD Inc* I am free to withdraw consent or discontinue participation at any time. I understand and agree to participate in the study conducted and videotaped by the *ASP.MD Inc*.

I understand and consent to the use and release of the videotape by *ASP.MD Inc*. I understand that the information and videotape 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 videotape and understand the videotape may be copied and used by *ASP.MD Inc* 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 with outside of *ASP.MD Inc* and *ASP.MD Inc's* client. 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 4: MODERATOR'S GUIDE USED IN THE STUDY

EHRUT Usability Test
Moderator's Guide

Administrator _____

Data Logger _____

Date _____ **Time** _____

Participant # _____

Location _____

Prior to testing

Confirm if user is available for uninterrupted session
Ask permission to record the study
Ensure EHRUT Platform is running properly

Prior to each participant:

Reset application
Start session recordings on Zoom

After each participant:

End session recordings *on Zoom*

After all testing

Back up all video and data files

Orientation (*X* minutes)

Thank you for participating in this study. Our session today will last ***15* minutes**. During that time you will take a look at an electronic health record system.

I will ask you to complete a few tasks using this system and 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. If you get lost or have difficulty, I want you to explore the app and find the solution on your own, as it will help us understand how easy or difficult accessing certain features is. 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.

The product you will be using today is *an Electronic health record management tool.*

We are recording the audio and screenshare of our session today. 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?

Preliminary Questions (*2-3* minutes)

May I know, what's your age?

What is your highest level of education?

What is your job title/appointment?

How long have you been working in this role?

How much experience do you have with computers/laptops in years?

Share the credentials with participants for login and ask them to share the screen. Once they login and share screen, the following tasks were given to the users.

Task 1: Looking up the patient and open the alerts section [Tests criterion 170.315(b)(11)]

Success:

- ☐ Easily completed
- ☐ Completed with difficulty or help:: Describe below
- ☐ Not completed

Comments:

Task Time:_____Seconds

Optimal Path: *Screen A* *Screen B* *Drop Down B*[1] *"OK" Button* *Screen X…*

- ☐ Correct
- ☐ Minor Deviations / Cycles :: Describe below
- ☐ Major Deviations :: Describe below

Comments:

Observed Errors and Verbalizations:

Comments:

Rating:

Overall, this task was: _____

Show participant written scale: "Very Difficult" (1) to "Very Easy" (5)

Administrator / Notetaker Comments:

Task 2: Checking the health report of the patient [Tests criterion 170.315(b)(11)]

Success:

- ☐ Easily completed
- ☐ Completed with difficulty or help:: Describe below
- ☐ Not completed

Comments:

Task Time:_____Seconds

Optimal Path: *Screen A → Screen B → Drop Down B[1] → "OK" Button → Screen X…*

- ☐ Correct
- ☐ Minor Deviations / Cycles :: Describe below
- ☐ Major Deviations :: Describe below

Comments:

Observed Errors and Verbalizations:

Comments:

Rating:

Overall, this task was: _____

Show participant written scale: "Very Difficult" (1) to "Very Easy" (5)

Administrator / Notetaker Comments:

Task 3: You have to override an alert by stating any of your preferred reason [Tests criterion 170.315(b)(11)]

Success:

- ☐ Easily completed
- ☐ Completed with difficulty or help :: Describe below
- ☐ Not completed

Comments:

Task Time:_____Seconds

Optimal Path: *Screen A* *Screen B* *Drop Down B[1]* *"OK" Button* *Screen X…*

- ☐ Correct
- ☐ Minor Deviations / Cycles :: Describe below
- ☐ Major Deviations :: Describe below

Comments:

Observed Errors and Verbalizations:

Comments:

Rating:

Overall, this task was: _____

Show participant written scale: "Very Difficult" (1) to "Very Easy" (5)

Administrator / Notetaker Comments:

Task 4: Checking the supporting documents on alert [Tests criterion 170.315(b)(11)]

Success:

- ☐ Easily completed
- ☐ Completed with difficulty or help :: Describe below
- ☐ Not completed

Comments:

Task Time:_____Seconds

Optimal Path: *Screen A → Screen B → Drop Down B[1] → "OK" Button → Screen X…*

- ☐ Correct
- ☐ Minor Deviations / Cycles :: Describe below
- ☐ Major Deviations :: Describe below

Comments:

Observed Errors and Verbalizations:

Comments:

Rating:

Overall, this task was: _____

Show participant written scale: "Very Difficult" (1) to "Very Easy" (5)

Administrator / Notetaker Comments:

Task 5: Adding feedback to an alert [Tests criterion 170.315(b)(11)]

Success:

- ☐ Easily completed
- ☐ Completed with difficulty or help :: Describe below
- ☐ Not completed

Comments:

Task Time:_____Seconds

Optimal Path: *Screen A → Screen B → Drop Down B[1] → "OK" Button → Screen X…*

- ☐ Correct
- ☐ Minor Deviations / Cycles :: Describe below
- ☐ Major Deviations :: Describe below

Comments:

Observed Errors and Verbalizations:

Comments:

Rating:
Overall, this task was: _____

Show participant written scale: "Very Difficult" (1) to "Very Easy" (5)

Administrator / Notetaker Comments:

Task 6: Checking the source information of on alert [Tests criterion 170.315(b)(11)]

Success:
- ☐ Easily completed
- ☐ Completed with difficulty or help :: Describe below
- ☐ Not completed

Comments:

Task Time:_____Seconds

Optimal Path: *Screen A → Screen B → Drop Down B[1] → "OK" Button → Screen X…*

- ☐ Correct
- ☐ Minor Deviations / Cycles :: Describe below
- ☐ Major Deviations :: Describe below

Comments:

Observed Errors and Verbalizations:

Comments:

Rating:
Overall, this task was: _____

Show participant written scale: "Very Difficult" (1) to "Very Easy" (5)

Administrator / Notetaker Comments:

Task 7: Running an alert feedback report [Tests criterion 170.315(b)(11)]

Success:
- ☐ Easily completed
- ☐ Completed with difficulty or help :: Describe below
- ☐ Not completed

Comments:

Task Time:_____Seconds

Optimal Path: *Screen A* *Screen B* *Drop Down B[1]* *"OK" Button* *Screen X…*

- ☐ Correct
- ☐ Minor Deviations / Cycles :: Describe below
- ☐ Major Deviations :: Describe below

Comments:

Observed Errors and Verbalizations:

Comments:

Rating:

Overall, this task was: _____

Show participant written scale: "Very Difficult" (1) to "Very Easy" (5)

Administrator / Notetaker Comments:

Final Questions (*X Minutes*)

How was your overall experience with the platform?

Did you find the tasks easy to complete?

Will you be able to perform these tasks again on your own?

Any suggestions to improve the experience of this platform?

Appendix 6: INCENTIVE RECEIPT AND ACKNOWLEDGMENT FORM

Acknowledgement of Receipt

I hereby acknowledge receipt of *$100* for my participation in a research study run by YellowSlice Pvt Ltd.

Printed Name:_____

Address:_____

_____

Signature:_____Date: _____

Usability Researcher:_____

Signature of Usability Researcher: _____

Date:_____

Witness: _____

Witness Signature:_____

Date:_____