



EHR Usability Test Report of MDLog

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

Version Tested: 5.0

Date of Usability Test: December 24th 2020 to February 2021

Date of Report: Dec 24th 2020

Last Updated: Mar 16th 2021

Report Prepared By: Nikhita Bhatlapenumarthy



Table of Contents

1. EXECUTIVE SUMMARY

2. INTRODUCTION

3. METHOD

3.1. PARTICIPANTS

3.2. STUDY DESIGN

3.3. TASKS

3.4. PROCEDURE

3.5. TEST LOCATION

3.6. TEST ENVIRONMENT

3.7. TEST FORMS AND TOOLS

3.8. PARTICIPANT INSTRUCTIONS

3.9. USABILITY METRICS

4. RESULTS

4.1. DATA ANALYSIS AND REPORTING

4.2. DISCUSSION OF THE FINDINGS

5. APPENDICES

5.1 APPENDIX 1: Sample Recruiting screener.

5.2 APPENDIX 2: Participant demographics

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

5.4 APPENDIX 4: Example Moderator's Guide

5.5 APPENDIX 5: System Usability Scale Questionnaire

5.6 APPENDIX 6: Incentive receipt and acknowledgment form



1. EXECUTIVE SUMMARY

A usability test of MDLog, version 5.0 was conducted between December 24th 2020 and February 2021. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in MDLog.

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

This study collected performance data on 5 tasks typically conducted on the EHR:

1. Demographic Data
2. CPOE - Medication order
3. CPOE – Laboratory order
4. Clinical decision support - Evidence-based decision support interventions
5. Implantable device

During the 60 minutes one-on-one usability test, each participant was greeted by the administrator and asked to review and sign an informed consent/release form; they were instructed that they could withdraw at any time. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using MDLog. 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 assistance in how to complete the task.

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 compensated with \$50 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 MDLog. Following is a summary of the performance and rating data collected on the MDLog.



Measure	N	No of Success Tasks	Task Success %	Task Success-Standard Deviation %	Path Deviation			Task Time(in Seconds)				Errors	Task Ratings (likert)	
					Actual Steps	Optimal Steps	Ratio	Mean	Optimal Time	Deviation (Observed Seconds)	Standard Deviation		Mean (SD)	Mean
Demographic Data	10	10	100%	0%	8	8	1	75	60	757	31.61	0 (0)	5	0
COPE – Laboratory order	10	10	100%	0%	6	6	1	38.4	30	384	20.59	0 (0)	4.9	0.31
Implantable device	10	10	100%	0%	6	6	1	37.3	30	373	19.21	0 (0)	4.9	0.31
COPE - Medication order	10	10	100%	0%	9	9	1	83.1	60	764	75.53	0 (0)	2.85	2.05
Clinical decision support - Evidence-based decision support interventions	10	10	100%	0%	11	11	1	286	210	2860	162.63	0 (0)	3.8	1.03

The results from the System Usability Scale scored the subjective satisfaction with the system based on performance with these tasks to be: 4-5.

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

- Major findings

The workflow around medication ordering is a bit challenging due to the multiple buttons and links displayed on the third party URL.

- Areas for improvement:

Reducing # of clicks and simplifying the interface for the following capabilities

- Medication Order
- Clinical Decision Support

2. INTRODUCTION

The EHRUT(s) tested for this study was MDLog, version 5.0. Designed to present medical information to healthcare providers in long term care settings, the MDLog consists of cloud based application whose web interface is accessible from wide range of devices ranging from PCs to Tablets including iPad. The usability testing attempted to represent realistic exercises and conditions.

In order to improve the usability of MDLog EHR, we follow the guidelines provided in NISITR 7741. Following the core of UCD that relies on systematic understanding of users and their environments, and iterative design and testing based on user performance objectives, we engage with our clients on a regular intervals through meeting and insist to have providers, NPs, Administrative staff and others who use our EHR regularly to be part of these meetings and try to get maximum on the field feedback from them.

We have integrated some third party applications into our system that can provide us with data analytics through which we review and enhance/optimize our workflows that would intern help us better the user experience.



We involve some key users in our Beta testing phase, to ensure we are delivering what the user need and would enhances the overall user experience .When new features or enhancements to current workflows are done, we make it a point to create graphical help animations which would help the users understand the changes in a faster and better way.

On a periodic basis we also conduct Usability testing, which will give us the calculated metrics of the usability of the application. Following the Agile methodology in our project life cycles, also gives us more opportunity to refine and address user issues much faster.

3. METHOD

3.1 PARTICIPANTS

A total of 10 participants were tested on the MDLog. Participants had no direct connection to the development of organization producing the MDLog. 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.

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.

Patient ID	Gender	Age Group	Education	Occupation	Professional Exp (in Months)	Computer Exp (in Months)	Product Exp (in Months)	Assistive Technology Needed
1	Female	50-59	Master’s Degree	Nurse Practitioner	168	360	12	No
2	Female	50-59	Master’s Degree	Nurse Practitioner	180	360	2	No
3	Female	50-59	Master’s Degree	Nurse Practitioner	360	240	9	No
4	Male	60-69	Master’s Degree	Office Manager	240	240	12	No
5	Female	30-39	Bachelor's Degree	Director of Operations	168	240	4	No
6	Female	50-59	Master’s Degree	Nurse Practitioner	60	240	6	No
7	Female	40-49	Master’s Degree	Nurse Practitioner	216	360	12	No
8	Female	50-59	Bachelor's Degree	Practice Manager	12	360	7	No
9	Female	30-39	Master’s Degree	Nurse Practitioner	36	360	3	No
10	Female	30-39	Master’s Degree	Nurse Practitioner	84	240	15	No



Total numbers of participants recruited are 10 matching the demographics in the section on participants.

Participants were scheduled for 60 minutes session and for each session administrator reset the systems to proper test conditions and briefed about test procedure. A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics as provided by the recruiting firm.

3.2 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 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 déviations
- Participants verbalisations (comment)
- Participant's satisfaction ratings of the system

3.3 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:

- Adding a new Medication order
- Adding a new Laboratory order
- Adding and updating Demographic Data
- Creating a Clinical decision support Alert
- Adding an Implantable device

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



3.4 PROCEDURES

Upon joining the meeting, participants were greeted; their identity was verified and matched with a name on the participant schedule. Each participant reviewed and signed an informed consent and release form shared the digital copy at the end of the test over an email.

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

For each task, the participants were given instructions on how to perform the task. Task timing began once the administrator finished giving the instructions. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Following the session, the administrator gave the participant the post-test questionnaire (e.g., the System Usability Scale), 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 indicating that they had received the compensation.

3.5 TEST LOCATION

The tests were conducted on a virtual platform using WebEx meetings. Based on the participant's availability, meetings were scheduled and invitations were sent in advance. Only the participant and administrator/Data Logger were in the call. To ensure that the environment was comfortable for users, noise levels were kept to a minimum.

3.6 TEST ENVIRONMENT

MDLog would typically be used in a Long Term Care Facility or Office Practice. The application is accessible over the internet.



For testing, the computer used was a PC laptop running Windows 10 Operating system. The participants used a mouse and a keyboard when interacting with MDLog.

The MDLog used an 13" LED screen of the PC laptop with resolution set to 1366*768. The application was set up by the MDOps according to the vendor's documentation describing the system set-up and preparation. The application is accessible over internet. 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).

3.7 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

3.8 PARTICIPANT INSTRUCTIONS

The administrator read the following instructions aloud to the each participant.

Thank you for participating in this study. Your input is very important. Our session today will last about 60 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 10 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 5 tasks to complete.

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

DATA SCORING

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

Measure	Rationale and Scoring
Effectiveness: Task Success	<p>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.</p> <p>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.</p> <p>Task times were recorded for successes.</p> <p>Observed task times divided by the optimal time for each task is a measure of optimal efficiency.</p> <p>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.</p> <p>Thus, if expert, optimal performance on a task</p>

	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.
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 5 (Very Easy) to 1 (Very Difficult). 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 MDLog 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

4. RESULTS

4.1 DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses

The usability testing results for the MDLog are detailed below. 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.

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

4.2 DISCUSSION OF THE FINDINGS

Measure	N	No of Success Tasks	Task Success %	Task Success-Standard Deviation %	Path Deviation			Task Time(in Seconds)				Errors	Task Ratings (likert)	
					Actual Steps	Optimal Steps	Ratio	Mean	Optimal Time	Deviation (Observed Seconds)	Standard Deviation	Mean (SD)	Mean	Standard Deviation
Demographic Data	10	10	100%	0%	8	8	1	75	60	757	31.61	0 (0)	5	0
COPE – Laboratory order	10	10	100%	0%	6	6	1	38.4	30	384	20.59	0 (0)	4.9	0.31
Implantable device	10	10	100%	0%	6	6	1	37.3	30	373	19.21	0 (0)	4.9	0.31
COPE - Medication order	10	10	100%	0%	9	9	1	83.1	60	764	75.53	0 (0)	2.85	2.05
Clinical decision support - Evidence-based decision support interventions	10	10	100%	0%	11	11	1	286	210	2860	162.63	0 (0)	3.8	1.03

EFFECTIVENESS



The subjects were able to complete 100% of the assigned tasks successfully

EFFICIENCY

100% of the tasks were completed by the subjects without any deviations from the optimal steps.

90% of those tasks were also completed within the mean time. For the remaining 10% of them which were the more complex Meaningful Use tasks like Clinical decision support system subjects required more time understanding the implementation conceptually and how to setup those capabilities. The feedback suggested that we rely more on Built-in Content like

- a. Default settings and Ready to use rules to trigger clinical decision system alerts

This approach will be more effective in improving the efficiency compared to customers doing the settings and defining rules for clinical decision system.

SATISFACTION

Subjects gave an average "Ease of Use" rating of 5 on a scale of 5 (very easy) to 1 (very difficult) and found the application more simpler to use. Overall they felt they could complete their daily tasks quickly without going through a long learning curve

MAJOR FINDINGS

Users were typically happy with the simplicity of the application when performing the daily tasks like Demographic updates, placing lab orders, adding implantable devices,

AREAS FOR IMPROVEMENT

1. Refine ordering medications interface to reduce number of clicks to bring it in line with rest of the application.
2. More Built-in Content particularly Rules to Trigger Clinical Decision support alerts