



---

# **IHS RESOURCE AND PATIENT MANAGEMENT SYSTEM SUMMATIVE USABILITY TESTING REPORT**

Version: 2.0

Date: 8/18/2021

Dates of Testing: 6/1/2021-6/11/2021

System Test Laboratory:

Contact Person: Jason Nakai, Information Architect/Usability Engineer, IHS Contractor

Email Address: Jason.Nakai@IHS.gov

Mailing Address: 5600 Fishers Lane, Rockville, MD 20857

## Version History

Date	Version	Revised By:	Revision/Change Description
7/30/2021	1.0	Jason Nakai	Initial Draft
8/18/2021	2.0	Jason Nakai	Updated

## Table of Contents

<b>1.0</b>	<b>Executive Summary .....</b>	<b>4</b>
1.1	Major Findings .....	5
1.2	Recommendations .....	6
<b>2.0</b>	<b>Introduction.....</b>	<b>7</b>
2.1	Purpose .....	7
2.2	Scope.....	7
<b>3.0</b>	<b>Method.....</b>	<b>8</b>
3.1	Roles and Responsibilities.....	9
3.1.1	Test Participants .....	10
3.2	Test Location .....	10
3.3	Test Environment.....	10
3.4	Test Tools .....	11
3.5	Tasks .....	11
3.6	Procedure .....	12
3.7	Usability Metrics.....	13
3.7.1	Data Scoring .....	13
<b>4.0</b>	<b>Results .....</b>	<b>15</b>
4.1	Data Analysis and Reporting .....	15
4.2	Discussion of Findings.....	15
4.2.1	Effectiveness .....	16
4.2.2	Efficiency .....	16
4.2.3	Satisfaction .....	16
4.2.4	Major Findings .....	17
4.2.5	Recommendations.....	18
<b>5.0</b>	<b>Acronym List.....</b>	<b>19</b>
<b>6.0</b>	<b>Appendix A: Participant and Test Result Data.....</b>	<b>20</b>

## 1.0 Executive Summary

A summative usability test (hereinafter referred to as “usability test”) of the Electronic Health Record (EHR) application was conducted during the months of May and June 2021 as part of the 21st Century Cures Act (21<sup>st</sup> CCA) (g)(3) Safety-Enhanced Design criterion. The purpose of this test was to evaluate and validate the usability of the current user interface, and provide evidence of user-centered design (UCD) practices in the application.

During the usability test, healthcare providers and other users matching the target demographic criteria participated in summative usability testing for each safety-enhanced design criterion and the associated capabilities.

This study collected performance data on the top tasks as identified by the owners of the criteria to be tested.

The criteria included in this test report are:

- (b)(3) ePrescribing (eRX)

During the approximately 60-minute one-on-one usability test sessions, each participant was greeted by the administrator who introduced the test. Participants were asked to share their prior EHR experience. The administrator logged in to the application and then passed control over to the participant to complete a series of tasks (given one at a time) using the application.

During each test session, the administrator timed the test and recorded user performance data. Participant screens and audio were also recorded for subsequent analysis.

The following types of data were collected for each participant:

- Demographic data
- Number of tasks successfully completed
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant’s verbalizations (comments)
- Participant’s satisfaction ratings of the system

All participant data was de-identified so that no correspondence could be made from the identity of the participant to the data collected.

The test method and metrics were based on the National Institute of Standards and Technology (NIST) Guide to the Processes Approach for Improving the Usability of Electronic Health Records (NISTIR 7741). Modifications were made where necessary to better evaluate the application against the contract goals and requirements. Following the conclusion of the test, participants were asked to complete a post-test questionnaire and were thanked for their participation.

The Task Satisfaction Rating is based on the following pre-defined of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants.

## 1.1 Major Findings

Based on the score of the Task Satisfaction Rating, the participants found the eRX component of the EHR easy to use.

Participants did state that the initial learning curve is steep and training is necessary. However, once they learned to use the application, participants completed tasks with great efficiency and effectiveness.

Most test participants felt the components were consistent and functioned as expected. The majority found the RPMS EHR to be an effective tool for completing their work tasks.

The top issues the test participants remarked on were:

- Font size and contrast made readability difficult
- Text was unable to be resized
- Button and menu text was not clear or intuitive
- More instructions especially for the functionality of the notes area that activates the action buttons (i.e. Approve, Accept, etc.)
- The right-click menus are not intuitive. Users did not know that they had to right-click to find the available actions

Tasks	Task Success	Task Satisfaction Rating (Scale 0-2)
	Mean %	% Rated 2 – Completed Easily
1. Create new prescription	100%	91%
2. Change prescription	91%	82%
3. Renew prescription	100%	82%
4. Cancel prescription	100%	100%

**Table 1: Criteria Success and Satisfaction Rating Summary**

## 1.2 Recommendations

Specific recommendations for the criteria are as follows:

- Default font size and contrast should be readable enough to meet Web Content Accessibility Guidelines (WCAG) 2.0 Level AA success criterion [1.4.3 Contrast \(Minimum\)](#)
- Text size should be able to be increased by the end user to a minimum of 200% to meet WCAG 2.0 Level AA success criterion [1.4.4 Resize text](#)
- Review all micro text to ensure that meaning and intent is clear
- Spell out acronyms
- Add clear and understandable instructions, hints or tool tips for complex or unintuitive actions. Examples of such actions include right-clicking on a change request to see the options available, and scrolling down to the bottom of a page to activate an approval button.

General recommendations for future development suggest that usability activities continue to be part of the development process for projects and/or products that involve user interfaces, and that usability lessons learned continue to be documented for potential future improvements.

## 2.0 Introduction

The Office of the National Coordinator for Health Information Technology (ONC) Health IT Certification Program is a voluntary certification program established by the Office of the National Coordinator for Health IT to provide for the certification of health IT.

The Indian Health Service (IHS) Office of Information Technology (OIT) has requested that the Resource and Patient Management System Electronic Health Record (RPMS EHR) achieve certification as part of the 21<sup>st</sup> CCA. As part of the certification criteria, (g)(3) Safety-Enhanced Design requires that summative usability testing be performed on specific criteria and the test data be provided as part of a final test report. The test report will follow the National Institute of Standards and Technology (NIST) Customized Common Industry Format Template for Electronic Health Record Usability Testing (NISTIR 7742).

Summative usability testing is a task-based evaluation that measures the ease of use of a completed product. The results are analyzed and compared to the usability requirements to determine if those requirements have been met.

### 2.1 Purpose

The purpose of this test was to evaluate and validate each safety-enhanced design criterion and the associated capabilities. The test ensures that the completed product meets the 21<sup>st</sup> CCA certification requirements concerning user-centered and safety-enhanced design.

### 2.2 Scope

The scope of usability testing is limited to testing user-involved tasks. Automated tasks or tasks without user interaction are not covered in this test. Functional testing is not covered in detail. Functionality is only tested as it pertains to the usability of the product or feature being tested.

The test was limited in scope to the following criterion:

- (b)(3) ePrescribing

## 3.0 Method

See Appendix A for Participant and Test data.

The test method and metrics were based on the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records (NISTIR 7741). Modifications were made where necessary to better evaluate the application against the contract goals and requirements.

The objective of this test was to uncover areas where the application performed well and areas where the application failed to meet the usability 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 capability and/or comparison with other EHR capabilities provided the same tasks are used. This testing serves as both a means to record or benchmark current usability and to identify areas where improvements must be made.

The application was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system (Task Satisfaction Rating)
  - 1(Very Difficult to 5 (Very Easy)

### 3.1 Roles and Responsibilities

Role/Function	Responsibilities
Project Manager/Criteria Owner	Responsible for the management, monitoring and tracking of the project and oversees all areas.
Usability Test Lead / Test Administrator	<ul style="list-style-type: none"> <li>• Ensures that usability testing is conducted successfully and meets all usability testing deadlines.</li> <li>• Provides application systems analysis for application testing activities.</li> <li>• Prepares required documentation at the program level for testing activities.</li> <li>• Monitors and escalates risks or concerns about achieving goals or meeting schedules to program leadership.</li> <li>• Prepares all testing instructions, scripts and materials for use in the testing session.</li> <li>• Performs analysis of testing results, prepares and delivers test report.</li> <li>• Moderates the test</li> <li>• Collects test data</li> </ul>
Test Observers	<ul style="list-style-type: none"> <li>• Provide any needed training or support</li> <li>• Monitor the testing session</li> </ul>
Test Participants	<ul style="list-style-type: none"> <li>• Complete the assigned tasks</li> <li>• Provide honest feedback on their experience</li> </ul>

**Table 3: Roles and Responsibilities**

### 3.1.1 Test Participants

There was a total of 11 test participants for this round of testing.

Participants in this test were:

- typical end-users such as physicians and medical providers
- trained to use the application prior to usability testing
- recruited by the 21<sup>st</sup> CCA project team and IHS criteria owners
- not compensated for participation
- had no direct connection to the development of the application
- given the same orientation and level of training as the actual end users
- assigned a participant ID initially based on scheduling order

Once participants were identified, they were scheduled for 60-minute one-on-one web conferencing (Skype, Adobe Connect) sessions. A calendar was used to keep track of the participants' schedule and a spreadsheet tracked participants' location (site) and contact information.

### 3.2 Test Location

The test was conducted remotely via the use of video conferencing and desktop sharing software (Microsoft Skype for Business, Adobe Connect).

### 3.3 Test Environment

The test participants were physically located at their normal duty stations, logged into their assigned workstations, and connected to the video conferencing software. The test administrator and observers were also physically distributed and connected via video conferencing software.

The test administrator shared his screen and was the only desktop visible during testing. Participants were given control of the test application through the test administrator's screen and used a mouse and keyboard when interacting with the application.

Technically, the system performance (i.e., response time) was not representative to what actual users would experience in a field implementation, as they were working through the test administrator's workstation and not their own.

### **3.4 Test Tools**

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

1. Demographic Questionnaire
2. Moderator's Guide
3. Post-test Questionnaire

The Moderator's Guide was devised so as to capture the required data.

Video conferencing software (MS Skype, Adobe Connect) was used to connect participants, the administrator and observers. This software was also used to record the video and audio of test sessions.

### **3.5 Tasks**

The testing scenarios and tasks were constructed to be realistic and representative of the kinds of activities a user would perform using the capabilities being tested. Tasks were chosen with the test objectives in mind to ensure that participants provided the most meaningful data possible. The tasks were arranged to facilitate a typical end-user workflow.

The testing tasks include:

1. Create new prescription
2. Change prescription
3. Renew prescription
4. Cancel prescription

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.

## 3.6 Procedure

Upon arrival, each participant was greeted by the administrator and matched to a name on the participant schedule. The participant was then assigned a participant ID.

The test administrator moderated the test session including administering instructions and tasks. The administrator also monitored task times, obtained post-task rating data, and took notes on participant comments.

Each participant was instructed to perform the tasks:

- 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.

Each participant used the same application version and was provided with the same set of instructions.

The administrator logged into the test environment and then instructed the user to request control. After log in, the user was instructed to complete a series of tasks (given one at a time) using the application. The participant was given a written copy of each task, and the administrator also read each task aloud and ensured the participant understood the task.

Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated that the task was successfully completed.

Scoring is discussed in [Section 3.7 Usability Metrics](#).

After completion of the testing tasks, the administrator gave the participant a post-test questionnaire (System Usability Scale), asked if they had any questions, and thanked them for their participation.

Each participant's demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire ratings were recorded into the participant spreadsheet.

Following each test session, the video recordings were reviewed and checked against the data logged in the participant spreadsheet. The participant spreadsheet was updated with any edits or additional information such as verbalizations.

### 3.7 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 by measuring participant success rates and errors
2. Efficiency by measuring the average task time and path deviations
3. Satisfaction by measuring task satisfaction ratings and SUS scores

#### 3.7.1 Data Scoring

The following table (Table 4) details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
<p><b>Effectiveness:</b> Task Success</p>	<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. 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. 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.</p>
<p><b>Effectiveness:</b> Task Failures</p>	<p>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 a “Failures.” No task times were taken for errors.</p> <p>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.<sup>11</sup> This should also be expressed as the mean number of failed tasks per participant.</p> <p>On a qualitative level, an enumeration of errors and error types should be collected.</p>

Measures	Rationale and Scoring
<p><b>Efficiency:</b> Task Deviations</p>	<p>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.</p>
<p><b>Satisfaction:</b> <b>Task Satisfaction Rating</b></p>	<p>Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the system 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."</p>

Table 4: Measure Scoring

## **4.0 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.

### **4.2 Discussion of Findings**

Based on the score of the Task Satisfaction Rating, the participants found the EHR easy to use. Participants did state that the initial learning curve is steep and training is necessary. However, once they learned to use the application, participants completed tasks with great efficiency and effectiveness.

The path taken to complete the tasks differed from participant to participant. This was influenced by the differing configuration of the test sites' EHR UIs. In spite of the varied paths to complete tasks, time per task was minimal and consistent, and errors were virtually non-existent.

All test participants felt the components were consistent and functioned as expected. The majority found the RPMS EHR to be an effective tool for completing their work tasks. Most said they would recommend this EHR to their colleagues.

The top issues the test participants remarked on were:

- Font size and contrast made readability difficult
- Text was unable to be resized
- Button and menu text was not clear or intuitive
- More instructions especially for the functionality of the notes area that activates the action buttons (i.e. Approve, Accept, etc.)
- The right-click menus are not intuitive. Users did not know that they had to right-click to find the available actions

### 4.2.1 Effectiveness

#	Tasks	# Participants	Task Success Rate - Mean %	Task Success Rate - Std Dev %	Task Errors Mean %	Task Errors Std Dev %
<b>(b)(3) ePrescribing</b>						
1	Create new prescription	11	100%	0%	0%	47%
2	Change prescription	11	91%	30%	0%	52%
3	Renew prescription	11	100%	0%	0%	50%
4	Cancel prescription	11	100%	0%	0%	0%

Table 5: Effectiveness

### 4.2.2 Efficiency

#	Tasks	Observed # Steps	Optimal # Steps	Task Time Observed Mean (seconds)	Task Time Std Dev (seconds)	Task Time Optimal (seconds)
<b>(b)(3) ePrescribing</b>						
1	Create new prescription	12	11	207	161	250
2	Change prescription	12	12	146	52	236
3	Renew prescription	14	14	128	42	180
4	Cancel prescription	4	4	46	15	90

Table 6: Efficiency

### 4.2.3 Satisfaction

#	Tasks	# Participants	Task Rating Likert Scale	Task Rating Mean	Task Rating Std Dev
<b>(b)(3) ePrescribing</b>					
1	Create new prescription	11	1-5	4.82	0.60
2	Change prescription	11	1-5	4.45	1.29
3	Renew prescription	11	1-5	4.64	0.81
4	Cancel prescription	11	1-5	5.00	0.00

Table 7: Satisfaction

#### 4.2.3.1 System Usability Scale (SUS)

The results from the System Usability Scale (SUS) from the post-test questionnaire, scored the subjective satisfaction with the system based on performance with the listed testing tasks by group.

System Usability Scale (SUS) Score	Score
(b)(3) ePrescribing	77.05

**Table 8: SUS Scores**

According to usability.gov, “[b]ased on research, a SUS score above a 68 would be considered above average and anything below 68 is below average”.

#### 4.2.4 Major Findings

Based on the score of the Task Satisfaction Rating, the participants found the eRx component of the EHR easy to use.

Participants did state that the initial learning curve is steep and training is necessary. However, once they learned to use the application, participants completed tasks with great efficiency and effectiveness.

Most test participants felt the components were consistent and functioned as expected. The majority found the RPMS EHR to be an effective tool for completing their work tasks.

The top issues the test participants remarked on were:

- Font size and contrast made readability difficult
- Text was unable to be resized
- Button and menu text was not clear or intuitive
- More instructions especially for the functionality of the notes area that activates the action buttons (i.e. Approve, Accept, etc.)
- The right-click menus are not intuitive. Users did not know that they had to right-click to find the available actions

#### 4.2.5 Recommendations

Overall recommendations focus on more effectively communication meaning to the end user, as well as enhancing readability. Specific recommendations for the criteria are as follows:

- Default font size and contrast should be readable enough to meet Web Content Accessibility Guidelines (WCAG) 2.0 Level AA success criterion [1.4.3 Contrast \(Minimum\)](#)
- Text size should be able to be increased by the end user to a minimum of 200% to meet WCAG 2.0 Level AA success criterion [1.4.4 Resize text](#)
- Review all micro text to ensure that meaning and intent is clear
- Spell out acronyms
- Add clear and understandable instructions, hints or tool tips for complex or unintuitive actions. Examples of such actions include right-clicking on a change request to see the options available, and scrolling down to the bottom of a page to activate an approval button.

## 5.0 Acronym List

Acronym	Description
CCA	21 <sup>st</sup> Century Cures Act
EHR	Electronic Health Record
eRX	ePrescribing
IHS	Indian Health Service
ISO	International Organization for Standardization
NIST	National Institute of Standards and Technology
OIT	Office of Information Technology
RPMS	Resource and Patient Management System
SESS	Software Engineering Support Services
UI	User Interface
WCAG	Web Content Accessibility Guidelines

Table 23: Acronyms

## 6.0 Appendix A: Participant and Test Result Data

Participant Identifier	Participant Gender	Participant Age	Participant Education	Participant Occupation/Role	Participant Professional Experience (months)	Participant Computer Experience (months)	Participant Product Experience (months)	Participant Assistive Technology Needs
TP1	Female	30-39	Pharm D	Clinical Informaticist	20	8	16	None
TP2	Female	40-49	Master's Degree	Clinical Informaticist	27	13	13	None
TP3	Female	40-49	Bachelor's Degree	Clinical Nurse Case Manager	30	18	9	None
TP5	Male	40-49	Doctorate, Master's Degree	Health Informaticist	30	13	13	None
TP8	Female	30-39	Doctorate	Pediatrician	25	5	2	None
TP9	Male	40-49	Pharm D	Health Systems Specialist	25	2	10	None
TP10	Male	50-59	Doctorate	Subject Matter Expert Physician	30	2	15	None
TP13	Male	50-59	AA Degree	IT Specialist, CAC	35	6	1	None
TP14	Female	30-39	Bachelor's Degree	Nurse Consultant	30	1	9	None
TP15	Male	40-49	Pharm D	Pharmacy Consultant, Clinical Informaticist	40	15	25	None
TP17	Male	70-79	Doctorate	Subject Matter Expert Physician	40	3	19	None

Task	Task Success Rate - Mean (%)	Task Success Rate - Standard Deviation (%)	Mean observed number of steps taken for the corresponding task	Optimal number of steps for the corresponding task
1. Create new prescription	100%	0%	12	11
2. Change prescription	91%	30%	12	12
3. Renew prescription	100%	0%	14	14
4. Cancel prescription	100%	0%	4	4

Task	Task Rating - Scale Type	Mean Task Rating (1-5)	Mean Task Rating Standard Deviation (1-5)
1. Create new prescription	Likert Scale	4.82	0.60
2. Change prescription	Likert Scale	4.45	1.29
3. Renew prescription	Likert Scale	4.64	0.81
4. Cancel prescription	Likert Scale	5.00	0.00

Task	Mean Task Time (seconds)	Standard Deviation for Task Time (seconds)	Observed Task Time (seconds)	Optimal Task Time (seconds)
1. Create new prescription	207	161	159	250
2. Change prescription	146	52	113	236
3. Renew prescription	128	42	95	180
4. Cancel prescription	46	15	42	90

## IHS Resource and Patient Management System

---

Task	Mean Task Errors (%)	Standard Deviation of Task Errors (%)
1. Create new prescription	27%	47%
2. Change prescription	45%	52%
3. Renew prescription	36%	50%
4. Cancel prescription	0%	0%