EHR Usability Test Report of VelociDoc v18.5

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

VelociDoc v18.5

Date of Usability Test: 11/07/2018

Date of Report: 11/16/2018

Report Prepared By: Practice Velocity, LLC

Adam Jennison
Director, Product
(815)708-4581
ajennison@practicevelocity.com
8777 Velocity Drive
Machesney Park, IL 61115

Table of Contents

PRACTICE VELOCITY®
Urgent Care Solutions®

8777 Velocity Drive | Machesney Park, Illinois 61115 | practicevelocity.com

EXECUTIVE SUMMARY

A usability test of VelociDoc v18.5 was conducted on 11/07/2018 by in-person. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). During the usability test,
16 healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative, tasks. This study collected performance data on 10 tasks typically conducted on an EHR:

1. Updating patient demographics
2. Creating and editing orders such as Laboratory Orders and Imaging Orders
3. Adding and editing a drug allergy in the patient record
4. Adding and editing a problem in the patient record
5. Recording the Universal Device Identifier code into the patient record
6. Creating the current medication list
7. Ordering, composing and editing new medications
8. Reviewing the alerts and contraindications to prescribed medications
9. Configuring a new clinical decision alert
10. Reconciling a patient's record with an external CCDA

During the 15-30 minute usability test, each participant was greeted by the administrator. All participants had prior experience with the EHR.

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 recorded user performance data on paper and electronically. The administrator did not give the participant assistance in how to complete the task.

A login with password and a "User Manual" consisting of step-by-step instructions and screenshots using a fictitious patient and representative data was provided to each participant, similar to the type of training material provided to any new user of the EHRUT.

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 from optimum path (noted but not counted)
- 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. Participants were not compensated for their time as testing was performed as part of the participants' workday.
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. See attached for a summary of the performance and rating data collected on the EHRUT.

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

Major findings

All participants were at least somewhat familiar with the VelociDoc v18.5 system. In most cases, users were able to work through the tasks quickly and without assistance. It was noted that there were some components of tasks that were confusing to the user. The administrator felt that most of the confusion lay in the fact that the participant did not perform that task in their regular workflow, as those who did performed well in the exercise.

Some users were effusively pleased with the addition of certain functionality, specifically the interface to import and reconcile C-CDA documents into the patient record.

In regards to deviations, the most difficult task was the electronic prescribing of medications. Users were generally unimpressed and frustrated with the outdated user interface. Otherwise, deviations were generally within expected ranges.

The administrators noted many small exclamatory comments, most not really relevant to the EHRUT, but to workflow. It was also noted that the participants occasionally did not take time to read the task fully and sometimes were forced to backtrack to reach the required conclusion. These were deemed to be deviations – not errors, as the desired conclusion was eventually reached.

Overall, administrators felt that participant satisfaction with the program and their performance using it was adequate-to-good. The majority of deviations from the optimal path were felt to be the result of user unfamiliarity with the task itself, not from within the program. Participants verified this to us, stating that they “never did this” and “this isn’t my job.”

Areas for improvement

More training and cross-training would make the system more familiar to all users. This would be primarily a function of the office management using materials provided by Practice Velocity, LLC.

Improvements could be made to the layout and usability of the e-prescribing module, to make the screens less busy and more streamlined.

Users totally unfamiliar with certain screens showed difficulty finding the appropriate places to enter certain data elements or search for criteria, expecting to find them elsewhere on the screen. These issues could be addressed to make those screens more intuitive.

INTRODUCTION

The EHR Under Test (EHRUT) tested for this study was VelociDoc v18.5. Designed to present medical information to healthcare providers in radiation oncology practices, the EHRUT is a secure web browser-based Electronic Health Record system. 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 EHRUT. To this end, measures of effectiveness, efficiency and user satisfaction such as time on task, ease of use and intuitiveness, were captured during the usability testing.

METHOD

PARTICIPANTS

A total of 10 participants were tested on the EHRUT. Participants in the test were medical providers and staff of an Urgent Care office. Participants were recruited from within the offices of current users of the EHR and were not compensated for their time. Participants had no direct connection to the development of or organization producing the EHRUT. Participants had the same orientation and level of training as the actual end users have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. Due to sampling bias given the small study, all participants happened to be female. The following is a table of participants by characteristics, including demographics, 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	Participant Computer Experience	Participant Product Experience	Participant Assistive Technology Needs
Phys1	Female	40-49	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Physician	132	264	9	No
Nurse1	Female	30-39	Bachelor's Degree	Nurse	84	240	9	No
Phys2	Female	40-49	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Physician	144	276	9	No
Nurse2	Female	30-39	Bachelor's Degree	Nurse	84	312	7	No
Rad1	Female	30-39	Bachelor's Degree	Radiology Tech	108	300	8	No
Phys3	Female	40-49	Doctorate degree (e.g., MD, DNP, DMD, PhD)	Physician	216	312	6	No
MA1	Female	20-29	Some college credit, no degree	Clinical Assistant	60	336	9	No
MA2	Female	20-29	Some college credit, no degree	Clinical Assistant	24	300	9	No
Rad2	Female	20-29	Bachelor's Degree	Radiology Tech	48	324	4	No
PCS1	Female	50-59	Bachelor's Degree	Patient Care Specialist	96	240	2	No

i.e., 10 participants were recruited and participated in the usability test. No participants failed to show for the study. Participants were scheduled for 30-45 minute sessions with at least 10 minutes in between each session for debrief by the administrator(s) to reset systems to proper test conditions.

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 participants reported improvements that could be made to the application to address better workflow and ease of use. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

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

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Task deviations from optimal path (noted but not counted)
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might perform with this EHR. Tasks were selected based on the twelve ONC CEHRT 2015 certification criteria, considering frequency of use, potential for risk to patient safety, and criticality of function. The Safety-Enhanced Design tasks for the twelve ONC CEHRT 2015 certification criteria included:

1. (a.1) CPOE –Meds
    - Record medication via CPOE
    - Change medication via CPOE
    - Display changed CPOE medication order
2. (a.2) CPOE – Labs
    - Record Lab order via CPOE
    - Change Lab order via CPOE
    - Display changed CPOE Lab order
3. (a.3) CPOE – Diagnostic Imaging
    - Record Imaging order via CPOE
    - Change Imaging order via CPOE
    - Display changed CPOE Imaging order
4. (a.4) Drug-drug, drug-allergy interaction checks for CPOE
    - Using CPOE, trigger a drug-drug interaction by entering a new medication order
    - Using CPOE, trigger a drug-allergy interaction by entering a new medication order
    - Adjust the severity level of a displayed drug-drug interaction
5. (a.5) Demographics
    - Record a patient's preferred language, date of birth, birth sex, race, ethnicity, sexual orientation, gender identity
    - Change the patient's preferred language, date of birth, birth sex, race, ethnicity, sexual orientation, gender identity

- Display the patient's changed preferred language, date of birth, birth sex, race, ethnicity, sexual orientation, gender identity

6. (a.6) Problem List
   - Record a problem to the problem list
   - Change a problem on the problem list
   - Display the active problem list
   - Display the historical problem list
7. (a.7) Medication list
   - Record a medication to the medication list
   - Change a medication on the medication list
   - Display the active medication list
   - Display the historical medication list
8. (a.8) Medication allergy list
   - Record a medication allergy
   - Change a medication allergy
   - Display the active medication allergy list
   - Display the historical medication allergy list
9. (a.9) Clinical Decision support
   - Add a CDS intervention and/or reference resource for each of the required elements
     - ❖ Problem list
     - ❖ Medication list
     - ❖ Medication Allergy List
     - ❖ At least one Demographic
     - ❖ Laboratory Test
     - ❖ Vital Signs
     - ❖ And a combination of at least 2 of the elements listed above
   - Trigger the CDS interventions/resources added using the applicable data elements from each of the required elements
   - View the intervention/resource information using the Infobutton standard for data elements in the problem list, medication list, and demographics
   - Trigger the CDS interventions/resources based on data elements in the problem list, medication list, and medication allergy list by incorporating patient information from a transition of care/referral summary
   - Access the following attributes for one of the triggered CDS interventions/resources: bibliographic citation, developer, funding source, release/revision date
10. (a.14) Implantable Device List
    - Record UDI
    - Change UDI Status
    - Access UDI, device description, identifiers, and attributes
11. (b.2) Clinical Information Reconciliation and Incorporation
    - Incorporate a CCDA and conduct reconciliation of the medications, medication allergies, and

problems in the CCDA with the information currently in the patient's record
    - Generate a new CCDA with reconciled data
12. (b.3) e-Prescribing
    - Create new prescription
    - Change prescription (dosage or duration)
    - Cancel prescription
    - Refill prescription
    - Receive fill status notification
    - Request and receive medication history information

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

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.

An informed consent and release form was not deemed necessary.

A Practice Velocity, LLC staff member administered this test. Because of the small size of the testing group, one administrator worked at the site administering instructions and tasks, monitoring task times, obtaining post-task rating data, and taking notes on participant comments. Additionally, he took notes on task success, path deviations, number and type of errors, and comments. The usability testing staff member was experienced with usability testing with significant background in user experience design with 1 year's experience with VelociDoc v18.5 specifically.

Participants were instructed to perform the tasks as quickly as possible making as few errors and deviations as possible.

All participant data was de-identified and kept confidential.

For each task, the participants were given a written copy of the task. The administrator as allowed to give immaterial guidance and clarification on tasks, but not instructions on use. Participants were provided with screenshots guiding them through each task for their use, if unable to perform the task without assistance.

Task timing began once the administrator finished reading the question. The task time was stopped once the participant had successfully completed the task. After each task, the participant was directed to enter a 0-5 score for Viewing, Entering and Editing each task into their individual scorecard. Each individual was thanked for their participation.

Following the session, the administrator distributed the post-test questionnaire (e.g., the Likert scale) to the participants.

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

TEST LOCATION

The test facility included a quiet testing room with tables and computers for the participants. 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 well-known to the participants.

TEST ENVIRONMENT

The EHRUT would be typically used in a healthcare office or facility. In this instance, the testing was conducted in a quiet office. For testing, the computers used were laptops running Windows 10. The participants used a keyboard and mouse when interacting with the EHRUT.

The VelociDoc v18.5 application was set up by the Practice Velocity, LLC staff. The application itself was running on a Windows platform using a test database accessed with an internet browser. 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).

PARTICIPANT INSTRUCTIONS

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

“The Usability Testing is a test of OUR system, not your performance. We welcome both positive and negative feedback on your experience navigating through the tasks that follow.

A specific task may or may not be included in your manual, according to your assigned role and your typical office workflow. Most of these tasks will already be familiar to you. Please complete them to the best of your ability. If you need help with a task, screenshots are provided to guide you through the process. If you use the screenshots, it is important to follow the numbered steps for each section in order to achieve the expected results. You should be aware that in many cases, problems and contraindications are meant to be generated. This is to test our system for use in real office situations.

A successfully completed task will be indicated by the expected appearance of the entry in the patient record.

To judge the system's ease of use, we ask you to record the time it takes to complete the tasks. When I say "BEGIN", please start your timer. When you have finished each task, please stop your timer.

If you find you are totally unable to complete the task and don't wish to continue with it, say aloud "QUIT".

At the end of each task, you will be asked to record your time and to score the task on a provided scorecard. Each task has 3 criteria to rank from 0-5:
Very Easy = 5; Easy = 4; Average = 3; Slightly Difficult = 2; Very Difficult = 1; Deficient = 0

Any score below 3 will trigger a return to the developer for adjustments/corrections. You will also have an opportunity to add comments about each section. These comments will be helpful to us in refining our system.

You have been assigned a specific patient in the system. Sign on using your assigned tester User Name and Password. Choose and open the patient assigned to you. The Patient's Initial visit has already been created in the system."

Participants were then given 12 tasks to complete.

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

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, was defined by taking a measure of optimal performance and multiplying by 2, thus allowing a time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was 100 seconds then allotted task time performance was 200 seconds. This ratio was 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 a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was observed. 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. Optimal paths were created when constructing the tasks. Because of the limited number of testing staff and participants it was determined that deviations would be observed and noted, but not counted.

Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until 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 review on the scoring form 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 EHRUT 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.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. No participants failed to follow session and task instructions resulting in their data excluded from the analyses.

The usability testing results for the EHRUT are attached.

The results from the Likert Scale scored the subjective satisfaction with the system based on performance with these tasks to be: 4.2. Broadly interpreted, scores under 3 represent systems with poor usability; scores over 4 would be considered above average.

Demographics entries posed no difficulties. New features of Demographics (170.315 a.5) regarding race and gender caused some comment, mostly regarding their actual use in a clinical setting. It was generally felt that the gender requirements could be awkward in actual use and would only be recorded if the patient self-identified.

Computerized Order and Imaging (170.315 a.2, 170.315 a.3), Medication Allergy (170.315 a.8) and Problem List (170.315 a.6) were fundamentally unchanged from previous versions and posed no difficulty to established users and only slight deviations from newer users and engendered little comment.

The Implantable Device entry was new to all users. It was agreed that there was a need to record the presence of the device, but potential entry of long UDI numbers was deemed to be burdensome.

The entering of current medications, Medication List (170.315 a.7) remained the same, as VelociDoc v18.5 utilizes the 3rd party provider, NewCrop. Using NewCrop, Electronic Prescribing (170.315 a.1, 170.315 b.3) and Drug-Drug Allergy Interaction (170.315 a.4) are combined into one operation and were tested together in Usability Testing. Testers were familiar with the interface but generally feel that it is poorly designed and needs a "facelift."

Clinical Decision Support (170.315 a.9) was familiar to those testing it. There was some appreciation expressed for the new alerts based on a combination of data elements.

The reconciliation of CCDAs (170.315 b.2) went well and was clearly understandable to all testers. This feature was a favorite among testers.

EFFECTIVENESS

For the most part, participants were assigned modules to test based on their real-life work duties. This meant that the areas tested were at least somewhat familiar to them and the concepts clear.

Participants generally felt that the system was effective in capturing the necessary data and that it was easy to use and relatively intuitive.

The administrator noted there was some difficulty on some screens in finding the correct button to click to get to a selection screen. The Implantable Device entry process was confusing to participants as they were a bit unclear on what needed to be entered. They wanted to enter the body part.

Verbal comments were mostly regarding the usefulness of a particular task in their own workflow.

Participants testing tasks with which they had no familiarity naturally had the most difficulty. There were several relatively new employees included in the 10 testers and their exposure to EHRs and VelociDoc v18.5 was limited. Generally, though, all the participants were able to perform the tasks easily and relatively quickly.

EFFICIENCY

The EHRUT was generally felt to be efficient in collecting all the needed data. Some participants felt that a few of the individual tasks were too complicated or confusing and could be streamlined to fewer steps.

SATISFACTION

Average satisfaction scores ranged from a low of 3.84 to a high of 4.73. On the individual scorecards, most tasks were given at least a "3" overall for their ease of viewing, entering and editing.

After testing, the post-test questionnaire was distributed. Each user was asked to rate their satisfaction with the EHRUT by responding to the questions with a number score from 1 (strongly disagree) to 5 (strongly agree). Ten users returned their questionnaires. The following is an overview of the scores returned in the questionnaire.

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

Overview of Final Questions

"What was your overall impression of this system?"

Responses mostly ranged from fair to good, noting it to be easy/user friendly. It was stated that the system was more intuitive than other EHRs they have used in the past.

"What aspects of the system did you like most?"

The ease of patient interface and the ability to import and reconcile CCDAs were the two aspects commented upon.

"What aspects of the system did you like least?"

A few users felt the system could use more integration between their other Healthcare applications. It was also mentioned that some participants would prefer use of the Chrome browser over Internet Explorer.

"Were there any features that you were surprised to see?"

There were no users who reported being surprised to see anything within the system.

"What features did you expect to encounter but did not see? That is, is there anything that is missing in this application?"

A few users commented on the desire for better in-app messaging.

"Compare this system to other systems you have used."

Most users compared the system favorably versus other EHRs. Two of these users had not used other systems.

"Would you recommend this system to your colleagues?"

Most users replied in the affirmative. Only one said no.

DISCUSSION OF RESULTS

MAJOR FINDINGS

All users were already familiar with the look and content of the EHRUT. Most of the participants found the interface to be easy to use and had no issues finding the screens and buttons they needed to follow the optimum path. Most difficulties/deviations arose because of individual user unfamiliarity with the specific type of task, in that they did not access those screens on a regular basis and weren't familiar with the exact placement of tabs, buttons, and fields or with specific terminology, codes, etc. Even with those difficulties, the correct path was fairly readily found and accomplished with success.

As expected, demographics and order entry tasks tested well. Although there were some changes to demographics in this version, the interface presented them in a familiar way and, other than discussion of the requirements themselves, no issues were encountered. Also testing well was the new requirement for entry of an Implantable Device. Users were most pleased with the interface to import and reconcile C-CDA records into a patient chart.

Those tasks involving the use of the NewCrop prescription pharmaceutical system showed few deviations but were performed slightly slower than anticipated. This may be due to the relatively low familiarity of the selected users or the NewCrop interface, itself. Even users familiar with the interface think it could be greatly improved with a new "look and feel."

Clinical Decision Support – the process to trigger the alerts was time consuming, but the process to configure them and check the bibliographic references was well-received.

Unexpectedly, the Problem List and Allergies showed a few more deviations from the optimal path than expected. However, these items still tested well and were performed in good time.

AREAS FOR IMPROVEMENT

The EHRUT provides all of the required elements for successful patient data management.

It is felt that most needed is more training and cross-training to make the system more familiar to all users. Of course, real-time office workflow may make additional training difficult.

Improvements could be made to the layout of some screens to make them less crowded and easier to find the correct field. Most of these comments pertained to the NewCrop interface.

There were some cases where the users showed difficulty understand what data needed to be entered in a field.

Users were very verbal about their desire for updates to the NewCrop screens.