Usability Test Report

Tecurologic LLC

PediNotes Version 5.1
EHR Usability Test Report of PediNotes Version 5.1


PediNotes Version 5.1

Date of Usability Test: 07/10/2017 - 07/13/2017
Date of Report: 07/24/2017
Report Prepared by: Tecurologic Test Laboratory
Steven B. Spedale, MD
President, Tecurologic
225-2146421
steve.spedale@tecurologic.com
PO Box 86536
Baton Rouge, LA, 70879

Table of Contents
1 EXECUTIVE SUMMARY 3
2 INTRODUCTION 6
3 METHOD 7
3.1 PARTICIPANTS 7
3.2 STUDY DESIGN 7
3.3 TASKS 8
3.4 PROCEDURE 9
3.5 TEST LOCATION 9
3.6 TEST ENVIRONMENT 10
3.7 TEST FORMS AND TOOLS 10
3.8 PARTICIPANT INSTRUCTIONS 10
3.9 USABILITY METRICS 11
4 RESULTS 12
4.1 DATA ANALYSIS AND REPORTING 12
4.2 DISCUSSION OF THE FINDINGS 15
5 APPENDICES 22
5.1 Appendix 1: Participant Overview 22
5.2 Appendix 2: Participant Demographics Summary 23
5.3 Appendix 3: Informed Consent Form 24
5.4 Appendix 4: Example Moderator’s Guide 25
5.5 Appendix 5: System Usability Scale Questionnaire 28
1. EXECUTIVE SUMMARY

A usability test of PediNotes version 5.1 modular EHR was conducted between 07/10/2017 – 07/13/2017 in Baton Rouge, LA by Tecurologic Test Laboratory. The purpose of this 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, 10 healthcare professionals matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

This study collected performance data on 31 tasks typically conducted in an EHR. For each of the tasks, the participant performed a function that would be categorized in following areas:

1. Computerized Provider Order Entry – Medications
2. Computerized Provider Order Entry – Laboratory
3. Computerized Provider Order Entry – Diagnostic Imaging
4. Drug-Drug, Drug-Allergy Interaction Checks for CPOE
5. Demographics
6. Problem List
7. Medication List
8. Medication Allergy List
9. Clinical Decision Support

During the 60-minute one-on-one usability test, each participant was greeted by the moderator and asked to review and sign an informed consent/release form; they were instructed that they could withdraw at any time. Participants had prior experience with some features but not all features of the EHR. Video training demonstrating how to use the EHR was given in advance for the participants to review. The moderator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the moderator and the data logger recorded the time data and user performance information on paper and electronically. The moderator did not give the participant assistance in how to complete the task unless they failed the task and required assistance on subsequent attempts.

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
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant’s verbalizations
- Participant’s satisfaction ratings of the system
  - 1 – Very Difficult
  - 2 – Difficult
  - 3 – Neither easy or difficult
  - 4 – Easy
  - 5 – Very Easy
- System Usability Scale (SUS)

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 all voluntary. 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.
The following is a summary of the performance and rating data collected on the EHRUT:

<table>
<thead>
<tr>
<th>Total Task Attempts</th>
<th>350</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Successful Task Attempts</td>
<td>310</td>
</tr>
<tr>
<td>Percent Success</td>
<td>89%</td>
</tr>
<tr>
<td>Optimal Steps per Task</td>
<td>14.75</td>
</tr>
<tr>
<td>Path Deviations (Observed/Optimal)</td>
<td>1.52</td>
</tr>
<tr>
<td>Optimal Task Time</td>
<td>89.93</td>
</tr>
<tr>
<td>Mean Task Time</td>
<td>90.97</td>
</tr>
<tr>
<td>Time Deviations (Observed/Optimal)</td>
<td>1.11</td>
</tr>
<tr>
<td>Total Errors</td>
<td>40</td>
</tr>
<tr>
<td>Mean Errors per Task</td>
<td>1.29</td>
</tr>
<tr>
<td>Mean Task Rating</td>
<td>4.34</td>
</tr>
<tr>
<td>SUS Rating</td>
<td>73.25</td>
</tr>
</tbody>
</table>

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

**MAJOR FINDINGS**

**Computerized Provider Order Entry – Medications**
- Participants struggled most with entering medications that were unfamiliar to them as they order medications for babies
- Several participants failed to correctly use medication modifiers
- Many participants ordered the incorrect drug due to similar drug names

**Computerized Provider Order Entry – Laboratory**
- Editing a laboratory was not straightforward to participants as an “edit” requires a delete/discontinue then reordering
- Several participants did not take advantage of search feature to find labs in existing order tree structure
- Participants did not understand what it meant to “access” the CPOE orders, which resulted in many extra steps/time

**Computerized Provider Order Entry – Diagnostic Imaging**
- Editing a diagnostic image was not straightforward to participants as an “edit” requires a delete/discontinue then reordering
- Several participants did not take advantage of search feature to find diagnostic images in existing order tree structure
- Participants did not understand what it meant to “access” the CPOE orders, which resulted in many extra steps/time

**Drug-Drug, Drug-Allergy Interaction Checks for CPOE**
- Participants expected a more obvious alert for a severe allergy reaction to medicine (different indication)
- Some participants did not realize the “View Reference” information was the mechanism to see the severity of interactions
Demographics
- The multiple-select race and ethnicity dialogs confused participants
  - Most didn’t understand that more than one item could be added at a time, as participants would add one item, save, then open the dialog again
  - The standard/detailed buttons confused participants (they asked for clarification)
  - Some didn’t immediately understand how to remove items that had been added
- Most participants did not understand the searching mechanism for races/ethnicities
- A few participants confused the Discharge/Transfer Date for the Death Date/Time on the discharge screen
- If the user wasn’t a physician, they had trouble finding the reason for death

Problem List
- Participants had difficult adding diagnoses via SNOMED concepts versus the current process of adding via ICD-10
- Few of the participants were familiar with the SNOMED concepts, therefore struggled with the selection of these diagnoses
- Participants had to associate the SNOMED concept with the ICD-10 diagnosis in order for other parts of the program to work, which was not an intuitive process

Medication List
- Participants struggled most with entering medications that were unfamiliar to them as they order medications for babies
- Several participants failed to correctly use medication modifiers to enter the medication as listed on the test sheet

Medication Allergy List
- All of the participants specialize in neonatal care therefore were unfamiliar with the process of adding allergies to patients and the Allergy screen in general
- The process of inactivating an allergy was not straightforward

Clinical Decision Support
- Participants did not understand how to use the InfoButton
- Participants did not realize they could hover over alerts to receive more information
- The process for reviewing the reference information was a deviation to the normal workflow for the participants

AREAS FOR IMPROVEMENT

Computerized Provider Order Entry – Medications
- Streamline the medication ordering process so that even non-familiar medications are presented in a way that clinicians understand
- Redesign the medication modifiers to be more clear

Computerized Provider Order Entry – Laboratory
- Provide an ability to “edit” a laboratory order without having to delete it
- Redesign the search window to make it more obvious that it is the first mechanism to find laboratory orders
- The method for accessing orders is different on the Orders screen from the Diagnosis/Care Plans screen (Diagnosis/Care Plans screen access could be more user friendly)

Computerized Provider Order Entry – Diagnostic Imaging
- Provide an ability to “edit” a diagnostic imaging order without having to delete it
- Redesign the search window to make it more obvious that it is the first mechanism to find diagnostic imaging orders
• The method for accessing orders is different on the Orders screen from the Diagnosis/Care Plans screen (Diagnosis/Care Plans screen access could be more user friendly)

**Drug-Drug, Drug-Allergy Interaction Checks for CPOE**
• Consider a redesign of the alert image for interactions that implies severity level
• Other alternatives for demonstrating severity level without clicking the “View Information” button are needed

**Demographics**
• Ability to add a standard versus detailed races/ethnicities needs to be more user friendly
• The method for removing races/ethnicities needs to be more straightforward for users, such as a traditional delete button
• The process for searching for races/ethnicities could be more intuitive
• Consider new layout for reason for death and death time to ensure appropriate data entry

**Problem List**
• Participants need education on SNOMED prior to usage of SNOMED diagnosis functionality
• Consider a redesign of the mapping process from SNOMED to ICD-10 to make more intuitive

**Medication List**
• Streamline the medication ordering process so that even non-familiar medications are presented in a way that clinicians understand
• Redesign the medication modifiers to be more clear

**Medication Allergy List**
• A mechanism to inactivate allergies is needed due to the confusion on process for inactivation
• Allergy screen can be simplified to make more intuitive for first time users

**Clinical Decision Support**
• Participants need education on the InfoButton mechanism for viewing context aware reference information
• In general, consider mechanisms to make the Clinical Decision Support reference information more effective so participants would use it
• More Clinical Decision Support items that result in better care for patients

2. **INTRODUCTION**
The EHRUT tested for this study was Tecurologic PediNotes 5.1. It has been designed to present medical information to healthcare providers in an inpatient hospital. The EHRUT consists of a Windows desktop WPF .Net Framework application connected to a database over a LAN. There is a Windows service component which sends and receives notifications/alerts to and from the EMR to other instances of the EMR. The EMR is used to enter all patient data, which includes (and is not limited to) the following:

- demographics
- family history
- relatives
- medication orders
- lab orders
- radiology orders
- lab results
- diagnoses
- allergies
- fluids
- vital signs
- blood orders
- pregnancy
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, deviations from optimal path, and task rating, were captured during the usability testing. The usability testing attempted to represent realistic exercises and conditions.

3. METHOD

3.1 PARTICIPANTS

A total of 10 participants were tested on the EHRUT(s). Participants in the test were Neonatology Nurse Practitioners and Physicians. Participants were recruited by Tecurologic. All participants volunteered at no charge. 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.

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, occupation/role and experience with EHRs. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

<table>
<thead>
<tr>
<th>Participant Identifier</th>
<th>Participant Gender</th>
<th>Participant Age</th>
<th>Participant Education</th>
<th>Participant Occupation/Role</th>
<th>Participant Computer Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>40-49</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>9y</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>40-49</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>15y</td>
</tr>
<tr>
<td>3</td>
<td>Female</td>
<td>30-39</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>11y</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>30-39</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>12y</td>
</tr>
<tr>
<td>5</td>
<td>Female</td>
<td>50-59</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>10y</td>
</tr>
<tr>
<td>6</td>
<td>Female</td>
<td>50-59</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>10y</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>40-49</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>15y</td>
</tr>
<tr>
<td>8</td>
<td>Female</td>
<td>40-49</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>6y</td>
</tr>
<tr>
<td>9</td>
<td>Female</td>
<td>40-49</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>19y</td>
</tr>
<tr>
<td>10</td>
<td>Female</td>
<td>50-59</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>20y</td>
</tr>
</tbody>
</table>

Ten participants (matching the demographics in the section on Participants) were recruited and ten participated in the usability test. No participants failed to show for the study.

Participants were scheduled for 60 minute sessions with 30 minutes in between each session for debrief by the moderator and data logger, 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.

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 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.

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:

<table>
<thead>
<tr>
<th>Task Number</th>
<th>Task ID</th>
<th>Task Category</th>
<th>Task Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A1.1</td>
<td>CPOE Medications</td>
<td>Record CPOE Medication Orders</td>
</tr>
<tr>
<td>2</td>
<td>A1.2</td>
<td>CPOE Medications</td>
<td>Change CPOE Medication Orders</td>
</tr>
<tr>
<td>3</td>
<td>A1.3</td>
<td>CPOE Medications</td>
<td>Access CPOE Medication Orders</td>
</tr>
<tr>
<td>4</td>
<td>A2.1</td>
<td>CPOE Laboratory</td>
<td>Record CPOE Laboratory Orders</td>
</tr>
<tr>
<td>5</td>
<td>A2.2</td>
<td>CPOE Laboratory</td>
<td>Change CPOE Laboratory Orders</td>
</tr>
<tr>
<td>6</td>
<td>A2.3</td>
<td>CPOE Laboratory</td>
<td>Access CPOE Laboratory Orders</td>
</tr>
<tr>
<td>7</td>
<td>A3.1</td>
<td>CPOE Imaging</td>
<td>Record CPOE Imaging Orders</td>
</tr>
<tr>
<td>8</td>
<td>A3.2</td>
<td>CPOE Imaging</td>
<td>Change CPOE Imaging Orders</td>
</tr>
<tr>
<td>9</td>
<td>A3.3</td>
<td>CPOE Imaging</td>
<td>Access CPOE Imaging Orders</td>
</tr>
<tr>
<td>10</td>
<td>A4.1</td>
<td>Drug-Drug/Drug-Allergy Interactions</td>
<td>Interventions: Drug-Drug</td>
</tr>
<tr>
<td>11</td>
<td>A4.2</td>
<td>Drug-Drug/Drug-Allergy Interactions</td>
<td>Interventions: Drug-Allergy</td>
</tr>
<tr>
<td>12</td>
<td>A4.3</td>
<td>Drug-Drug/Drug-Allergy Interactions</td>
<td>Adjustments (Drug-Drug Severity Levels)</td>
</tr>
<tr>
<td>13</td>
<td>A5.1</td>
<td>Demographics</td>
<td>Record Demographics</td>
</tr>
<tr>
<td>14</td>
<td>A5.2</td>
<td>Demographics</td>
<td>Change Demographics</td>
</tr>
<tr>
<td>15</td>
<td>A5.3</td>
<td>Demographics</td>
<td>Access Demographics</td>
</tr>
<tr>
<td>16</td>
<td>A6.1</td>
<td>Problem List</td>
<td>Record Problem List</td>
</tr>
<tr>
<td>17</td>
<td>A6.2</td>
<td>Problem List</td>
<td>Change Problem List</td>
</tr>
<tr>
<td>18</td>
<td>A6.3</td>
<td>Problem List</td>
<td>Access Active Problem List</td>
</tr>
</tbody>
</table>
Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Constructing appropriate tasks is of critical importance to the validity of a usability test. These are the actual functions, but most tasks contain larger and more fleshed out context that aligns with the sample data sets available in the tested EHR.

3.4 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. Each participant reviewed and signed an informed consent and release form. A representative from the test team witnessed the participant’s signature. To ensure that the test ran smoothly, two staff members participated in this test, the usability administrator and the data logger. The testing staff had between 5-20 years of experience with EHR software design and usability, with some having backgrounds in medicine (one MD) and several with backgrounds in computer science (BS and MS in Computer Science).

The moderator of the session administered the instructions and tasks. The moderator 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 a written copy of the task. Task timing began once the moderator finished reading the task prompt. 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 moderator gave the participant the post-test questionnaire, the System Usability Scale, 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.

3.5 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, the administrator, and the data logger 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.

3.6 TEST ENVIRONMENT
The EHRUT would be typically be used in a healthcare inpatient facility. In this instance, the testing was conducted in offices of Tecurologic. For testing, the computer used a windows desktop running Microsoft Windows 10. The participants used a keyboard and mouse when interacting with the EHRUT. The EHRUT used a 30” display with a resolution of 1440x900 landscape. The application was set up by Tecurologic according to the vendor’s documentation describing the system set-up and preparation. The application itself was running on a .NET framework platform using a training 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).

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

The Moderator’s Guide was devised so as to be able to capture required data. The participant’s interaction with the EHRUT was captured and recorded digitally with screen capture software running on the test machine. A Camtasia screen recording was created for each participant, and verbal comments were recorded with a microphone. The test session was electronically transmitted to a network server. The sessions were then viewed again by the development team to gather additional information for test analysis.

3.8 PARTICIPANT INSTRUCTIONS
The moderator reads the following instructions aloud to each participant:

Thank you for participating in this study. Your input is very important. Our session today will last about 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 is it would be useful to you, and how we could improve it. 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.

Please consider the following as your complete your testing:

- This is to test functionality only
These tests are not intended to be clinical
The goal is to test functionality of PediNotes, not the clinical actions of the user
For drug ordering, if a specific dosage is not specified, choose any
For drug ordering, if an indication is not specified, choose any
The diagnosis standard used for this testing is SNOMED CT. If a SNOMED CT code does not have a direct mapping to ICD-10 code, choose any ICD-10 code

After following the procedural instructions, participants were shown the EHR and as their first assignment, 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 “Go.” 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 31 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 Tecurologic PediNotes 5.1 by measuring participant success rates and errors
2. Efficiency of Tecurologic PediNotes 5.1 by measuring the average task time and path deviations
3. Satisfaction with Tecurologic PediNotes 5.1 by measuring ease of use ratings

DATA SCORING

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

<table>
<thead>
<tr>
<th>Measures</th>
<th>Rationale and Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectiveness:</td>
<td>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.</td>
</tr>
<tr>
<td>Task Success</td>
<td>The total number of successes were calculated for each task and then divided by the total number of times that task was attempted.</td>
</tr>
<tr>
<td></td>
<td>Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency.</td>
</tr>
<tr>
<td></td>
<td>Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times was operationally defined by taking multiple measures of optimal performance and multiplying by a factor of 1.25, which allows some time buffer because the participants are presumably not trained to expert performance.</td>
</tr>
</tbody>
</table>
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 “Failures.” No task times were taken for errors.

The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant.

On a qualitative level, an enumeration of errors and error types would 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.

The task deviations would be reported. Optimal paths (i.e., procedural steps) would be recorded when constructing tasks.

Efficiency:
Task Time
Each task was timed from when the administrator said “Begin” until the participant said, “Done.” If he or she failed to say “Done,” the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.

Satisfaction:
Task Rating
Participant’s subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate “Overall, this task was:” on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants.

Common convention is that average ratings for systems judged easy to use should be 3.3 or above.

To measure participants’ confidence in and likeability of PediNotes 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.”

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. There were a number of new features in the software that the users were unfamiliar with. Not only were they unaware of the features presence in the software, they were unfamiliar with the terminology and purpose of the functions. Therefore, an explanation of these items to the tester was required after failure of the task. The usability testing results for the EHRUT are detailed below.
| Task ID | Task Category | Task Description | Task Success - Mean (%) | Observed # Steps | Optimal # Steps | Deviations (Observed / Optimal) | Mean (seconds) | Standard Deviation (seconds) | Optimal Time | Deviations (Observed / Optimal) | Task Errors - Mean (%) | Standard Deviation (%) | Task Rating | Task Rating - Standard Deviation |
|---------|---------------|------------------|-------------------------|-----------------|----------------|-----------------------------|----------------|-------------------------------|--------------|-----------------------------|-----------------------|--------------------------|-------------|---------------------------------
<p>| A1.1 | CPOE Medications | Record CPOE Medication Orders | 100 | 64.0 | 50 | 2.0 | 45.0 | 19.0 | 4.0 | 0.0 | 0.0 | 3.6 | 0.8 |
| A1.2 | CPOE Medications | Change CPOE Medication Orders | 90.9 | 34 | 22 | 1.5 | 149.0 | 50.0 | 140.0 | 1.1 | 9.1 | 28.6 | 4.1 | 0.6 |
| A1.3 | CPOE Medications | Access CPOE Medication Orders | 100 | 6 | 1 | 6.0 | 45.0 | 32.0 | 28.0 | 1.6 | 0.0 | 0.0 | 4.3 | 0.7 |
| A2.1 | CPOE Laboratory | Record CPOE Laboratory Orders | 100.0 | 11.0 | 7.0 | 1.6 | 67.0 | 16.0 | 63.0 | 1.1 | 0.0 | 0.0 | 4.4 | 0.5 |
| A2.2 | CPOE Laboratory | Change CPOE Laboratory Orders | 76.9 | 15.0 | 10.0 | 1.5 | 52.0 | 23.0 | 56.0 | 0.9 | 23.1 | 40.7 | 4.4 | 0.7 |
| A2.3 | CPOE Laboratory | Access CPOE Laboratory Orders | 100.0 | 3.0 | 2.0 | 1.5 | 7.0 | 10.0 | 6.0 | 1.2 | 0.0 | 0.0 | 4.6 | 0.5 |
| A3.1 | CPOE Imaging | Record CPOE Imaging Orders | 100.0 | 13.0 | 7.0 | 1.9 | 37.0 | 16.0 | 38.0 | 1.0 | 0.0 | 0.0 | 4.6 | 0.5 |
| A3.2 | CPOE Imaging | Change CPOE Imaging Orders | 83.3 | 11.0 | 10.0 | 1.1 | 34.0 | 12.0 | 26.0 | 1.3 | 16.7 | 36.6 | 4.7 | 0.5 |
| A3.3 | CPOE Imaging | Access CPOE Imaging Orders | 100.0 | 1.0 | 1.0 | 1.0 | 3.0 | 5.0 | 8.0 | 0.4 | 0.0 | 0.0 | 4.7 | 0.5 |
| A4.1 | Drug/Drug-Allergy Interactions | Interventions: Drug-Drug | 71.4 | 34.0 | 29.0 | 1.2 | 157.0 | 34.0 | 143.0 | 1.1 | 28.6 | 42.5 | 4.0 | 0.8 |
| A4.2 | Drug/Drug-Allergy Interactions | Interventions: Drug-Allergy | 100.0 | 13.0 | 13.0 | 1.0 | 48.0 | 14.0 | 48.0 | 1.0 | 0.0 | 0.0 | 4.4 | 0.5 |
| A4.3 | Drug/Drug-Allergy Interactions | Adjustments (Drug-Drug Severity Levels) | 76.9 | 30.0 | 29.0 | 1.0 | 130.0 | 68.0 | 140.0 | 0.9 | 23.1 | 40.7 | 4.2 | 0.8 |
| A5.1 | Demographics | Record Demographics | 83.3 | 86.0 | 80.0 | 1.1 | 426.0 | 100.0 | 385.0 | 1.1 | 16.7 | 36.6 | 4.1 | 0.9 |
| A5.2 | Demographics | Change Demographics | 71.4 | 83.0 | 80.0 | 1.0 | 380.0 | 124.0 | 390.0 | 1.0 | 28.6 | 42.5 | 4.2 | 0.8 |
| A5.3 | Demographics | Access Demographics | 100.0 | 10.0 | 7.0 | 1.4 | 90.0 | 39.0 | 135.0 | 0.7 | 0.0 | 0.0 | 4.4 | 0.7 |
| A6.1 | Problem List | Record Problem List | 100.0 | 13.0 | 13.0 | 1.0 | 69.0 | 16.0 | 56.0 | 1.2 | 0.0 | 0.0 | 4.7 | 0.5 |
| A6.2 | Problem List | Change Problem List | 100.0 | 13.0 | 13.0 | 1.0 | 74.0 | 23.0 | 75.0 | 1.0 | 0.0 | 0.0 | 4.5 | 0.7 |
| A6.3 | Problem List | Access Active Problem List | 90.9 | 3.0 | 2.0 | 1.5 | 23.0 | 18.0 | 19.0 | 1.2 | 9.1 | 28.6 | 4.6 | 0.5 |
| A6.4 | Problem List | Access Problem List History | 100.0 | 2.0 | 2.0 | 1.0 | 12.0 | 11.0 | 13.0 | 0.9 | 0.0 | 0.0 | 4.6 | 0.5 |
| A7.1 | Medication List | Record Medication List | 100.0 | 54.0 | 40.0 | 1.4 | 244.0 | 51.0 | 215.0 | 1.1 | 0.0 | 0.0 | 4.0 | 0.9 |
| A7.2 | Medication List | Change Medication List | 90.9 | 25.0 | 25.0 | 1.0 | 96.0 | 38.0 | 116.0 | 0.8 | 9.1 | 28.6 | 3.6 | 1.6 |
| A7.3 | Medication List | Access Active Medication List | 76.9 | 5.0 | 2.0 | 2.5 | 35.0 | 34.0 | 19.0 | 1.8 | 23.1 | 40.7 | 3.7 | 1.6 |</p>
<table>
<thead>
<tr>
<th></th>
<th>Medication List</th>
<th>Access Medication List History</th>
<th>90.9</th>
<th>4.0</th>
<th>3.0</th>
<th>1.3</th>
<th>23.0</th>
<th>19.0</th>
<th>13.0</th>
<th>1.8</th>
<th>9.1</th>
<th>28.6</th>
<th>3.9</th>
<th>1.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>A7.4</td>
<td>Medication List</td>
<td>Access Medication List History</td>
<td>100.0</td>
<td>9.0</td>
<td>10.0</td>
<td>0.9</td>
<td>48.0</td>
<td>19.0</td>
<td>56.0</td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
<td>4.6</td>
<td>0.7</td>
</tr>
<tr>
<td>A8.1</td>
<td>Medication Allergy List</td>
<td>Record Medication Allergy List</td>
<td>90.9</td>
<td>18.0</td>
<td>14.0</td>
<td>1.3</td>
<td>87.0</td>
<td>47.0</td>
<td>60.0</td>
<td>1.5</td>
<td>9.1</td>
<td>28.6</td>
<td>4.4</td>
<td>0.7</td>
</tr>
<tr>
<td>A8.2</td>
<td>Medication Allergy List</td>
<td>Change Medication Allergy List</td>
<td>83.3</td>
<td>3.0</td>
<td>2.0</td>
<td>1.5</td>
<td>12.0</td>
<td>5.0</td>
<td>18.0</td>
<td>0.7</td>
<td>16.7</td>
<td>36.6</td>
<td>4.7</td>
<td>0.5</td>
</tr>
<tr>
<td>A8.3</td>
<td>Medication Allergy List</td>
<td>Access Active Medication Allergy List</td>
<td>100.0</td>
<td>3.0</td>
<td>2.0</td>
<td>1.5</td>
<td>9.0</td>
<td>5.0</td>
<td>19.0</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
<td>4.6</td>
<td>0.7</td>
</tr>
<tr>
<td>A8.4</td>
<td>Access Medication Allergy List History</td>
<td>Access Medication Allergy List History</td>
<td>83.3</td>
<td>6.0</td>
<td>6.0</td>
<td>1.0</td>
<td>44.0</td>
<td>26.0</td>
<td>41.0</td>
<td>1.1</td>
<td>16.7</td>
<td>36.6</td>
<td>4.7</td>
<td>0.7</td>
</tr>
<tr>
<td>A9.1</td>
<td>Clinical Decision Support</td>
<td>CDS Configuration</td>
<td>66.7</td>
<td>6.0</td>
<td>3.0</td>
<td>2.0</td>
<td>85.0</td>
<td>41.0</td>
<td>85.0</td>
<td>1.3</td>
<td>33.3</td>
<td>43.0</td>
<td>4.3</td>
<td>0.8</td>
</tr>
<tr>
<td>A9.2</td>
<td>Clinical Decision Support</td>
<td>CDS Intervention Interaction</td>
<td>59.9</td>
<td>4.0</td>
<td>1.0</td>
<td>4.0</td>
<td>26.0</td>
<td>29.0</td>
<td>3.0</td>
<td>9.0</td>
<td>2.9</td>
<td>3.1</td>
<td>4.3</td>
<td>0.3</td>
</tr>
<tr>
<td>A9.3</td>
<td>Clinical Decision Support</td>
<td>Source Attributes</td>
<td>83.3</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>6.0</td>
<td>8.0</td>
<td>0.5</td>
<td>16.7</td>
<td>36.6</td>
<td>4.5</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>A9.4</td>
<td>Clinical Decision Support</td>
<td>Source Attributes - Drug-Drug, Drug Allergy Interaction Checks</td>
<td>83.3</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>4.0</td>
<td>6.0</td>
<td>8.0</td>
<td>0.5</td>
<td>16.7</td>
<td>36.6</td>
<td>4.5</td>
<td>0.9</td>
</tr>
</tbody>
</table>

**Total:**

89.26 | 18.72 | 14.79 | 1.55 | 90.97 | 31.06 | 84.81 | 1.12 | 10.05 | 19.90 | 4.34 | 0.76
The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be: 73.25 (see below), with 1 = Strongly Disagree, 2 = Disagree, 3 = Neither Agree or Disagree, 4 = Agree, 5 = Strongly Agree.

<table>
<thead>
<tr>
<th>Participant</th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
<th>#9</th>
<th>#10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I think that I would like to use this system frequently</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. I found the system unnecessarily complex</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3. I thought the system was easy to use</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>4. I think that I would need the support of a technical person to be able to use this system</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. I found the various functions in this system were well integrated</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6. I thought there was too much inconsistency in this system</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7. I would imagine that most people would learn to use this system very quickly</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>8. I found the system very cumbersome to use</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>9. I felt very confident using the system</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>10. I needed to learn a lot of things before I could get going with this system</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Score</td>
<td>90</td>
<td>87.5</td>
<td>70</td>
<td>67.5</td>
<td>67.5</td>
<td>72.5</td>
<td>80</td>
<td>70</td>
<td>67.5</td>
<td>60</td>
</tr>
</tbody>
</table>

4.2 DISCUSSION OF THE FINDINGS

Of the 9 categories of tasks, participants did very well with CPOE-Medications, CPOE – Laboratory, CPOE – Diagnostic Imaging, Problem List, and Medication Allergy List tasks. Participants struggled with Drug-Drug/Drug-Allergy Interactions, Demographics, Medication List, and Clinical Decision Support tasks. There were specific tasks within each area of testing that were problematic either due to participant unfamiliarity with the tasks or participant uncertainty in requirements of the tasks themselves. While several of the specific features being tested were new to participants, the areas that participants performed well were similar to the processes currently used in their current workflow. For instance, CPOE of medications, laboratory, or diagnostic imaging was part of the clinician’s normal workflow, so the ordering process was not difficult. However, the problems occurred with ordering items that participants were not familiar with.

EFFECTIVENESS

Overall, the EHR was effective for the tasks performed. Of the 31 tasks, 14 were completed with no failures, while 24 were completed with 2 errors or less. There were 350 total attempts at the 31 tasks with 310 successes. That average success rate for the 31 tasks combined is 89%. Problem list was the most successful criteria with only 1 error in the 41 task attempts (97.6% success). The Clinical Decision Support tasks had an average of 77.6%, which was the lowest of any of the criteria tested and had 12 total errors. Demographics tasks were also problematic, with an average of 84.9% success rate. These 2 areas should be evaluated critically for more effective usability.

Another area examined was the areas that failed and the types of errors that caused failure. The errors were broken up into 3 categories:

1. User not knowing how to use function in the program
2. User not knowing where the function exists in the program
3. User error in actions or data input

Of the 40 task failures:

1. 8 errors were for not knowing how to use a function in the program
2. 12 errors were for not knowing where the function exists in the program
3. 20 errors were for user error in actions or data input
The errors were somewhat split between the 3 areas which could indicate both improved design as well as education on the changes. Training should be provided for the new functions to allow the users to understand how to use the features before a live clinical scenario.

EFFICIENCY
The efficiency of the program was broken into two major areas:
1. Path Deviation
2. Task Time

Path Deviation
There were a high number of path deviations for some tasks with little or no deviations for others. For instance, Clinical Decision Support had large deviations compared to other areas of testing. There was an average of 2 steps for every optimal step. Clinical Decision Support task 3 had 4 extra steps for every optimal step, indicating users were very unfamiliar with the functionality.

In addition, there were several tasks within certain criteria that had specific deviations that were high. CPOE Medications task 3 had a 6 ratio of observed deviations compared to the optimal path, indicating there was significant confusion on the task. Users were not clear on the task expectations which resulted in extra clicking, switching between pages, among other deviations from the optimal path. The average steps for most of the tasks was higher than the optimal path. This indicates that the optimal path was not clear to the users and there could be better ways of performing the tasks than the participants were completing.

The most average additional steps were taken on the following criteria:
1. A1.1 - Record CPOE Medication Orders - 14 additional steps
2. A1.2 - Change CPOE Medication Orders - 12 additional steps
3. A7.1 - Record Medication List – 14 additional steps

Task Time
In many cases, the task time reflected the deviations indicated previously. There were some cases where users spent a long period of time looking for a function but without performing unnecessary steps, but this was not common. There were several cases where the average time taken was less than the optimal time to perform. This could be because of the familiarity with the functions in the program after using repetitively. There were few functions where the time to perform the task was significantly higher than the optimal task.

The most additional time was taken on the following criteria:
1. A1.1 - Record CPOE Medication Orders – 85 seconds additional time
2. A5.1 - Record Demographics – 41 seconds additional time
3. A7.1 - Record Medication List – 29 seconds additional time
4. A8.2 - Change Medication Allergy List – 27 seconds additional time

There were also several tasks where participants performed quicker than the optimal time. This indicates that the PediNotes design was effective for performing the task.

SATISFACTION
The tasks ratings were a good indication of how each of the tasks were viewed by the participants. The average rating for all 31 tasks combined was 4.34, would indicate that PediNotes is above the expected rating for easy to use EHR's (3.3). All of the functions had ratings above 3.3, which was a good indication the users were satisfied with how the program functioned. Although there was great success in using the
Medication List functions in the program (in terms of successes and failures), most users rated this function lower than the other functions, with an average rating of 3.95. This was the only task category that averaged below 4. There should be changes to the medication program to make it easier to use.

The SUS average score for PediNotes was 73.25, which is above the common convention “above average score” of 68 for usability testing. 6 of the 10 participants scored the EHR above average. The comments from these participants indicates they were uncertain how to use some of the newer features of the program and there should be changes made to make it easier to use. These comments and suggestions will be valuable going forward.

MAJOR FINDINGS

Computerized Provider Order Entry – Medications
There was a common issue among participants with CPOE medications that was reflected in all medication tasks. Participants struggled with entering medications that were unfamiliar to them as they order medications for babies. Many participants ordered the incorrect drug due to similar drug names. The CPOE medication tasks were designed to mirror the testing steps for ONC certification. In future tests, the medication related tasks will be modified to reflect the normal workflow of neonatologists/neonatal nurse practitioners. However, the feedback from the testing should be evaluated as an opportunity to adapt the medication ordering more easy for non-neonatal practitioners. Several participants failed to correctly use medication modifiers, which indicates that the modifier entry was not straightforward. The medication order dialog in general has many more options than lab/imaging orders, so there was a large opportunity for user error.

Computerized Provider Order Entry – Laboratory
Laboratory CPOE was straightforward for the participants aside of editing labs. Editing a laboratory was not clear to participants as an "edit" requires a delete/discontinue then reordering. This was a common complaint during testing. Several participants did not take advantage of search feature to find labs in existing order tree structure. Given the labs were not located in the order tree where the participants would expect them, they spent unnecessary time trying located the items. The search function would have saved them time.

Computerized Provider Order Entry – Diagnostic Imaging
Diagnostic imaging CPOE was straightforward for the participants aside of editing imaging orders. Editing a diagnostic image was not clear to participants as an “edit” requires a delete/discontinue then reordering. Similar to the laboratory ordering, several participants did not take advantage of search feature to find diagnostic images in existing order tree structure.

Drug-Drug, Drug-Allergy Interaction Checks for CPOE
Drug-Drug, Drug-Allergy interaction checks provided much feedback that the Tecurologic development team had not previously received. Participants did not realize the “View Reference” information was the mechanism to see the severity of interactions, which ultimately led to the discussion on alert severity levels. Participants expected a more obvious alert for a severe allergy reaction to medicine (different indication). They expected a distinction between a severe and moderate level alert, consistent with other alerts in PediNotes.

Demographics
The ONC required changes to race and ethnicity selection was introduced in PediNotes 5.1 with varied feedback. The users of PediNotes previously selected from a small number of races/ethnicities which made the selection process straightforward. With the introduction of requirements to provide an ability to choose from all races/ethnicities in both 170.207(f)(1) and 170.207(f)(2), the PediNotes mechanisms to select these values was difficult for users. Most participants did not understand the searching mechanism for races/ethnicities. The ability to select multiple race and ethnicities confused participants. PediNotes uses
Admission/Discharge/Transfer (ADT) information from hospitals to automatically populate Race/Ethnicity information, so the entry of race/ethnicity was a deviation from the normal workflow. There were specific areas in the race/ethnicity dialogs that should be reviewed:

- Most didn’t understand that more than one item could be added at a time, as participants would add one item, save, then open the dialog again
- The standard/detailed buttons confused participants (they asked for clarification)
- Some didn’t immediately understand how to remove items that had been added

In addition to race/ethnicity selection, several nurse practitioner participants confused the Discharge/Transfer Date for the Death Date/Time on the discharge screen. Also, if the user wasn’t a physician, they had trouble locating the reason for death. The specific errors related to clinical role was an unexpected finding from the testing.

**Problem List**
As with previous User Centered Design testing, participants had difficult adding diagnoses via SNOMED concepts versus the current process of adding via ICD-10. Few of the participants were even aware of the SNOMED coding system in general, therefore struggled with the selection of these diagnoses. Participants had to associate the SNOMED concept with the ICD-10 diagnosis in order for other parts of the program to work, which was not an intuitive process. Outside of the SNOMED selection, there were few problems with this criterion.

**Medication List**
As mentioned in CPOE – Medication finding section, participants struggled most with entering medications that were unfamiliar to them as they order medications for babies. The use of the medication list to verify active medications versus the full medication history is a common function for clinicians, therefore had few issues overall.

**Medication Allergy List**
All of the participants specialize in neonatal care therefore were unfamiliar with the process of adding allergies to patients. When babies are born, they do not have allergies, therefore the Allergies screen is seldom used within the users of PediNotes. Nevertheless, the allergy entry was intuitive aside from inactivating allergies. Users often deleted the allergy rather than inactivating it, similar to laboratory/imaging ordering, which would indicate a redesign was necessary.

**Clinical Decision Support**
The introduction of the InfoButton with Clinical Decision Support was an area of concern. Participants did not understand how to use the InfoButton. Many participants did not realize they could hover over the alert description to receive more information about the alert. This may be resolved with education in the future. We also received feedback that the review of reference information was a deviation to the normal workflow for the participants. In most situations, they would not receive a Clinical Decision Support and click for details on the item. In general, they would have an awareness of the background on a Clinical Decision Support before the alert was provided.

**AREAS FOR IMPROVEMENT**
The usability testing was very helpful for the development team. It demonstrated the variations of each of the participants in completing the same tasks not only in duration but the number of steps taken to accomplish the tasks. Some participants took much longer than others to accomplish the tasks, but this did not necessarily indicate they performed the tasks with more deviations. The areas that one participant struggled often was an area that all participants struggled. There are many areas for improvement in the areas tested.

**Computerized Provider Order Entry – Medications**
Tecurologic should streamline the medication ordering process so that even non-familiar medications are presented in a way that clinicians understand. Medication CPOE should be straightforward regardless of the
specialty of the clinician. A redesign the medication modifiers is needed as well. In general the medication order dialog should be reordered based on feedback from the participants to make ordering with modifiers more intuitive.

**Computerized Provider Order Entry – Laboratory**
PediNotes should contain the ability to “edit” a laboratory order without having to delete it. This was a common issue for both laboratory orders and imaging orders. An emphasis on the search window is needed to make it more obvious that it is the first mechanism to find laboratory orders. The method for accessing orders is different on the Orders screen from the Diagnosis/Care Plans screen (Diagnosis/Care Plans screen access could be more user friendly). Participants did not understand what it meant to “access” the CPOE orders, which resulted in many extra steps/time. In future tests, this will be more clear on the proctor sheets.

**Computerized Provider Order Entry – Diagnostic Imaging**
PediNotes should contain the ability to “edit” a laboratory order without having to delete it. This was a common issue for both laboratory orders and imaging orders. An emphasis on the search window is needed to make it more obvious that it is the first mechanism to find diagnostic imaging orders. The method for accessing orders is different on the Orders screen from the Diagnosis/Care Plans screen (Diagnosis/Care Plans screen access could be more user friendly). Participants did not understand what it meant to “access” the CPOE orders, which resulted in many extra steps/time. In future tests, this will be more clear on the proctor sheets.

**Drug-Drug, Drug-Allergy Interaction Checks for CPOE**
It was evident from the testing that the interaction checking alert should give an indication of the severity of the alert without the user being required to click to determine this. Tecurologic should consider a redesign of the alert image for interactions that implies severity level. Other alternatives for demonstrating severity level without clicking the “View Information” button are needed as well.

**Demographics**
This was one of the most difficult areas for users. PediNotes should provide an ability to add a standard versus detailed races/ethnicities that is more user friendly. The process for searching for races/ethnicities could be more intuitive. The ability to add from the full listing of races/ethnicities needs to be isolated to only when the user clearly wants to enter a specific race/ethnicity, but allow them to enter from the standard races with ease. The method for removing races/ethnicities needs to be more straightforward for users, such as a traditional delete button. Click to remove function should be removed from other locations within PediNotes as the problems were evident. Also, consider new layout for reason for death and death time to ensure appropriate data entry. Nurse practitioner workflow resulted in some errors entering this information, therefore, the placement of the information should fit more within their existing processes (preparing a discharge note for the attending).

**Problem List**
Participants will need education on SNOMED prior to usage of SNOMED diagnosis functionality in the future. A help document/training within the SNOMED functions would be particularly effective. A redesign of the mapping process from SNOMED to ICD-10 should be considered so users are more aware of why they have to undergo this process. Tooltip instructions could make this function more intuitive.

**Medication List**
As mentioned in the CPOE medications areas for improvement, the medication ordering process should be streamlined so that even non-familiar medications are presented in a way that clinicians understand. The modifiers should be reorganized so that users see/use them in the way that are recommended for specific dosages.

**Medication Allergy List**
Allergy entry could be simplified although it is seldom used within neonatology. Allergy screen in general could be simplified to make more intuitive for first time users. A mechanism to inactivate allergies is needed due to the confusion on process for inactivation. Perhaps an “inactivate” button is the solution for straightforward allergy inactivation.
Clinical Decision Support
The feedback on Clinical Decision Support was unexpected. Participants stated that they were unaware of the hover details as well as the right click option to view more information. PediNotes needs to redesign these items so it is clear that more information is present than initially visible. There should be education/training on the information contained within each Clinical Decision Support. Participants need education on the InfoButton mechanism for viewing context aware reference information. In general, consider mechanisms to make the Clinical Decision Support reference information more effective so participants would use it. Finally, more Clinical Decision Support items that result in better care for patients, seek more opportunities for adding them within application.

Overall, the testing was very effective in gaining insight in the usability of PediNotes. Once the areas for improvement have been completed, most testing and training with the features would be ideal to prepare for use in live clinical scenarios.
APPENDICES

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

- Appendix 1: Participant Overview
- Appendix 2: Participant Demographics Summary
- Appendix 3: Informed Consent Form
- Appendix 4: Example Moderator’s Guide
- Appendix 5: System Usability Scale Questionnaire
Appendix 1: Participant Overview

1. 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 […]? [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. 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
1. What is your current position and title? (Must be healthcare provider)
   • □ RN: Specialty ______________________
   • □ Physician: Specialty _________________
   • □ Resident: Specialty __________________
   • □ Administrative Staff
   • □ Other [Terminate]

2. 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
1. How many years have you used an electronic health record?
2. How many EHRs do you use or are you familiar with?
# Appendix 2: Participant Demographics Summary

<table>
<thead>
<tr>
<th>Participant Identifier</th>
<th>Participant Gender</th>
<th>Participant Age</th>
<th>Participant Education</th>
<th>Participant Occupation/Role</th>
<th>Participant Professional Experience</th>
<th>Participant Computer Experience</th>
<th>Participant Product Experience</th>
<th>Participant Assistive Technology Needs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>40-49</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>9y Neonatology</td>
<td>108</td>
<td>67</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>40-49</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>15 y Neonatology</td>
<td>150</td>
<td>67</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Female</td>
<td>30-39</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>6y 2m Neonatology</td>
<td>132</td>
<td>43</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>30-39</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>12y Neonatology</td>
<td>144</td>
<td>11</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>Female</td>
<td>50-59</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>27y Neonatology</td>
<td>120</td>
<td>37</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>Female</td>
<td>50-59</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>20y Neonatology</td>
<td>120</td>
<td>67</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>40-49</td>
<td>Doctorate degree</td>
<td>Physician</td>
<td>8y Neonatology</td>
<td>180</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>Female</td>
<td>40-49</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>19y 6m Neonatology</td>
<td>72</td>
<td>67</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>Female</td>
<td>40-49</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>21y Neonatology</td>
<td>228</td>
<td>67</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>Female</td>
<td>50-59</td>
<td>Master's Degree</td>
<td>NNP</td>
<td>30y Neonatology</td>
<td>240</td>
<td>67</td>
<td>No</td>
</tr>
</tbody>
</table>
Appendix 3: Informed Consent

Informed Consent

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

Agreement
I understand and agree that as a voluntary participant in the present study conducted by Tecurologic LLC 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 Tecurologic LLC.
I understand and consent to the use and release of the videotape by Tecurologic LLC. 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 Tecurologic LLC 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 Tecurologic LLC and Tecurologic LLC’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 Moderators Guide

Only 1 task is presented here for illustration.

170.315 (a)(6) - Problem List

Date: __________  Name: ________________________________

During this test you will be asked to:
- Record, change, and access a patient’s active problem list.

Pre Loaded Data:
Inpatient: Health IT Developer selects a patient and pre-loads the following problems test data for:

- Hospital Day #1 (2 days before test date):
  Essential Hypertension (disorder); SNOMED code: 59621000

- Hospital Day #2 (1 day before test date):
  Diabetes Mellitus Type 2 (disorder); SNOMED code: 44054006

Task #1 – Record Problem List

- Select Patient: 170.315.a6
- Show Problem list
- Add following diagnoses with current date as onset
  - Acquired Hypothyroidism (disorder); SNOMED code: 111566002
  - Chronic Rejection of Renal Transplant (disorder); SNOMED code: 236578006

Reviewer

Ask user to complete user rating

SCORING

Task #1 – Record Problem List

Optimal Path:
- Select Patient: 170.315.a6
- Navigate to Diagnosis/Care Plan
- Demonstrate Diagnoses
- Navigate to Add Diagnosis
• Enter Snomed Code or Diagnosis Name Acquired Hypothyroidism (disorder): SNOMED code: 111566002
• Select Save
• Navigate Select Add Diagnosis
• Enter Chronic Rejection of Renal Transplant (disorder); SNOMED code: 236578006
• Save

Task Times:
Expected time: ________
Actual time: ________

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

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

Observed Errors and Verbalizations:
Comments:
Final Questions *(X Minutes)*

What was your overall impression of this system?
What aspects of the system did you like most?
What aspects of the system did you like least?
Were there any features that you were surprised to see?
What features did you expect to encounter but did not see? That is, is there anything that is missing in this application?
Compare this system to other systems you have used. Would you recommend this system to your colleagues?
Appendix 5: System Usability Scale Questionnaire

Strongly Agree = 5, Agree = 4, Neutral = 3, Disagree = 2, Strongly disagree = 1

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