Improving the User Experience with Discount Site-Specific User Testing

Robert P. Pierce[1]   Bernie R. Eskridge[1]   Brandi Ross[2]   Margaret A. Day[1]   Brooke Dean[1]   Jeffery L. Belden[1]

[1] University of Missouri Health Care, Columbia, Missouri, United States
[2] Tiger Institute, Cerner Corporation, Columbia, Missouri, United States

Address for correspondence Margaret A. Day, MD, MSPH, Department of Family and Community Medicine, University of Missouri, One Hospital Drive, Columbia, MO 65212, United States (e-mail: daym@health.missouri.edu).

Abstract

Objectives   Poor electronic health record (EHR) usability is associated with patient safety concerns, user dissatisfaction, and provider burnout. EHR certification requires vendors to perform user testing. However, there are no such requirements for site-specific implementations. Health care organizations customize EHR implementations, potentially introducing usability problems. Site-specific usability evaluations may help to identify these concerns, and "discount" usability methods afford health systems a means of doing so even without dedicated usability specialists. This report characterizes a site-specific discount user testing program launched at an academic medical center. We describe lessons learned and highlight three of the EHR features in detail to demonstrate the impact of testing on implementation decisions and on users.

Methods   Thirteen new EHR features which had already undergone heuristic evaluation and iterative design were evaluated over the course of three user test events. Each event included five to six users. Participants used think aloud technique. Measures of user efficiency, effectiveness, and satisfaction were collected. Usability concerns were characterized by the type of usability heuristic violated and by correctability.

Results   Usability concerns occurred at a rate of 2.5 per feature tested. Seventy percent of the usability concerns were deemed correctable prior to implementation. The first highlighted feature was moved to production despite low single ease question (SEQ) scores which may have predicted its subsequent withdrawal from production based on post implementation feedback. Another feature was rebuilt based on usability findings, and a new version was retested and moved to production. A third feature highlights an easily correctable usability concern identified in user testing. Quantitative usability metrics generally reinforced qualitative findings.

Conclusion   Simplified user testing with a limited number of participants identifies correctable usability concerns, even after heuristic evaluation. Our discount usability approach to site-specific usability has a role in implementations and may improve the usability of the EHR for the end user.

Keywords
► electronic health records
► user-centered design
► user-computer interface

Background and Significance

The federal incentives of the Health Information Technology for Economic and Clinical Health Act of 2009 fostered rapid adoption of electronic health records (EHRs) in the United States over the past decade. By 2017, 96% of hospitals and 80% of physician offices reported using a certified EHR.[1] Despite widespread adoption, the usability of the EHR remains a serious concern. Usability is defined as the extent to which technology helps users "achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use."[2] Poor usability of EHRs contributes to provider dissatisfaction and burnout while increasing risks to patients.[3–5]

To improve EHR usability and enhance the safety and efficiency of EHR technology, the Office of the National Coordinator for Health Information Technology (ONC) established vendor requirements for user-centered design (UCD) practices and summative user testing in the 2014 Edition Certification Criteria for EHR Technology.[6] These requirements, termed "safety-enhanced design" criteria, applied to core EHR areas including computerized provider order entry (CPOE), drug interaction checking, medication and allergy list management, clinical decision support, electronic prescribing, and clinical information reconciliation. The 2015 Edition safety-enhanced design criteria extended these core areas to also include problem list management, demographics, and implantable device lists.[7] Therefore vendors certifying EHR technology in these core areas must show that UCD processes have been applied and summative testing has been performed.

The EHR tested by the vendor for ONC certification may be quite different than the one encountered by the end user. Health care organizations may customize their systems based on practice design and risk management and compliance concerns, such that workflows in these core areas differ substantially from those tested and certified by the vendor.[8] Such configuration changes may significantly impact usability and alter the validity of any previous user testing.[9] Health care organizations make many configuration decisions outside of the core EHR areas for which UCD processes and testing are required of vendors. For example, the vendor may certify to safety-enhanced design for CPOE, but the details related to each particular order – the various ways the order is accessed, the number and order of additional data fields required, and the alerting related to that order – may be organizational decisions. Providers may feel that EHR systems should be adapted to individual workflow preferences,[10] resulting in pressure on the organization from providers to further customize workflows. Any of these site-specific customizations may impact the usability of the EHR technology.

The difference between the usability of a system in certification testing and post-implementation encountered by the user has been termed the "EHR usability reality gap."[11] This gap persists, in part, because health care systems often do not have usability specialists or resources available to assist with configuration of their site-specific implementa-tions. "Discount usability" refers to fast and inexpensive usability methods which forego expensive, elaborate metrics in favor of earlier feedback and iteration.[12,13] Discount usability methods may afford health care systems lacking usability resources a way to leverage efficient usability methods to improve their users' experience.

Objectives

In this report, we describe the initial results from the user testing component of a discount site-specific usability program developed at the University of Missouri. We describe the program and characterize the effectiveness of pre-implementation user testing of features considered for the health care organization's specific EHR implementation. We describe the number and types of usability concerns identified. We describe three of the features in detail to highlight the impact of testing on implementation decisions and on user experience.

Methods

University of Missouri (MU) Health Care includes six hospitals and over 50 outpatient clinics, serving 26 counties in central Missouri. MU Health Care has historically employed a UCD process incorporating user input and multidisciplinary teams and using methods including heuristic evaluation and iterative design. This process and these methods have been applied to EHR features new to the implementation of the Cerner Millennium EHR platform at MU Health Care, as permitted by feature scope, resources, and completion time-lines. But historically the UCD methods did not include a user testing program. With the goal of optimizing usability of newly configured EHR features, a team of two physicians, two physician informaticists, a nurse informaticist, and a solution architect developed a simplified user testing program to supplement the existing UCD methods.

We attempted to usability test every implementation change which significantly altered nurse or provider workflows. We also attempted to test any feature which introduced workflows new to our implementation. Features were excluded from testing if they were trivial (for example, renaming an order, the addition of a dropdown selection to an existing dialog, or the extension of an existing feature to a new user role). Features were also excluded when the implementation was urgent and time did not permit testing or the feature was not expected to impact many users. User testing was clustered into three separate 2-week long test events. Each test event targeted approximately five users, with one test session per user. Each test event was separately evaluated by the health care system's institutional review board (IRB) who made the determination that the projects were quality improvement activities and not human subject research, and therefore did not require additional IRB review.

Tasks and Scenarios

One team member (R.P.P.) with prior experience performing summative usability testing for vendor EHR certification designed the user tasks and scenarios. As much as possible,

tasks were patient- and problem-centered. The scenarios included the test patient name and some clinical context. Multiple test patients were used, but in order to minimize "down time" spent opening and closing charts during a test session, the features and tasks were configured using as few test patients as possible. The tasks involved either evoking or interacting with the feature under test. For interruptive alerts, the tasks were designed to evoke the alert; the participants were asked to work through the alert. For passive alerts, test patients were configured with the passive alerts displayed, and participants were given tasks which involved identification of and interaction with the passive alerts. In the case of clinical pathways, the tasks involved using the information in the scenario, interacting with the clinical decision support in the pathway to develop the most appropriate plan in each test patient.

Tasks for each test event were aggregated in a moderator guide and distributed to team members for review and revision. The test patients and scenarios were configured in one or more non-production test domains. Each task and scenario was then evaluated with at least one "dry run" using a volunteer provider. The tasks were revised a final time and a final moderator guide distributed to each team member. Each feature was evaluated with one to eight tasks. Thirteen features were evaluated over the course of three test events (►Table 1).

Recruitment

We recruited six participants per test event. Participants were recruited with email from MU Health Care leadership requesting assistance with usability testing of new feature functionality. We targeted a convenience sample of attending or resident physicians or physician assistants for each test event. In the first test event, recruitment emails were sent to all medical resident, fellow, and faculty users. Because the second test event included two features related to the care of children, the second test event targeted only pediatrics and family medicine providers. Recruitment for the first test event was difficult. Therefore, participants in subsequent test events were offered $50 monetary compensation for their participation.

Test Protocol

Informed consent was obtained, modeled on that in the Customized Common Industry Format Template for EHR Usability Testing (NISTIR 7742).[14] One hour was allotted for each test. Participant audio and screen actions were recorded using Morae software (TechSmith, Okemos, Michigan, United States) in the first test event. It provided detailed aggregate summative test data and includes metrics such as mouse movement, mouse clicks, and task time, and it integrated survey instruments such as the single ease question (SEQ).[15] However, Morae is no longer sold or supported.[15] We recorded audio and screen actions in subsequent test events using the cloud-based software platform Zoom (zoom.us, San Jose, California, United States). Testing was performed remotely with one moderator and one observer. Participants logged into the test domain(s) with their own usernames and passwords so that the EHR configurations

used in testing were the same as they used in daily practice. Participants were given no instruction regarding speed or accuracy but were invited to use concurrent think-aloud protocol. Retrospective probing was used after each task. Testing generally took the entire scheduled hour. Two features were evaluated only qualitatively. For 11 features, summative testing was performed and at least one task was evaluated quantitatively. User satisfaction was assessed using the SEQ[16] after each task. Throughout the process, including both task development and testing, the notable issues with each feature were recorded by the team members.

Analysis

One author (R.P.P.) reviewed all recordings. For the tasks evaluated quantitatively, he extracted or computed task time and errors by type, including mistakes, slips, UI errors, and scenario errors. The same author also characterized each user task as completed with ease (no workflow deviations, interruptions or errors), completed with difficulty (completed but with workflow deviations, interruptions, or errors), or failed to complete. If a feature was found not to work according to its functional design, it was labeled a "build defect" and fixed by the solution architect team member as soon in the testing event as possible. All other issues identified were categorized as "usability concerns." The rate of build defects and usability concerns per feature was computed by dividing the total number of build defects or usability concerns over the three test events by the number of features evaluated. Usability concerns were reviewed by three of the authors (R.P.P., M.A.D., and J.L.B.) and classified as one or more violations of Nielsen's 10 general principles for interaction design.[17] Concerns were also reviewed by three of the authors familiar with the MU Health Care implementation (B.R.E., B.R., and R.P.P.) and characterized as to their correctability. Discordant classifications were discussed by the reviewing authors and consensus reached. Usability concerns were reviewed among the test group and with the users and health care organization committees responsible for governance of the relevant domains. Usability test reports were submitted to health care organization leadership.

Results

Seventeen unique providers participated in user testing. Ten participants (59%) were female; nine (53%) were residents or fellows, one participant was a physician assistant, and the remainder were attendings. Participants had an average of 7.1 years in the profession (range 0.25–33) and an average of 5.3 years of experience with the EHR (range 0.25–18). The demographics and specialties of the participants are shown in ►Table 2.

We identified four build defects over the course of three test events, a rate of 0.3 build defects per feature tested. Examples of build defects include a missing order option in a clinical pathway and failure to flag a needed control as required in a user interface. All build defects were corrected before moving the feature into production. Thirty-three usability concerns

Table 1 Features, scenarios, and tasks used in testing

Feature	Description	Testing strategy	Tasks
Test event #1			
Interruptive alert: low value care	Based on the Choosing Wisely program; alerts user on entry of orders for cervical cytology smears, low back imaging, and carotid ultrasound when available documentation does not support performing the study.		Scenario: H. Z-test. Patient is on the schedule for a pap smear. She has no history of abnormal paps. 1. Order a pap smear.* Scenario: A. Z-test. Patient has a history of carotid stenosis and is due for a carotid ultrasound. 2. Order a carotid ultrasound.* Scenario: F. Z-test. The patient complains of 10 days of low back pain without radiation, no red flag symptoms. 3. Order lumbar spine X-rays.*
Passive alert: annual screens	Alerts provider when depression screening, alcohol misuse screening, or fall risk is overdue or is positive; alerts provider when there has been no provision of information of advance directive information		Scenario: C. Z-test is in the office for a visit. 1. Which annual assessment screens are overdue? 2. Which annual assessment screens are positive? Scenario: To address the high BMI, you counsel the patient about diet and exercise 3. Document your intervention. Scenario:K. Z-test is in the office for a visit. 4. Which annual assessment screens are due? 5. Which annual assessment screens are positive? Scenario: To address the positive alcohol/drug screen, you counsel the patient about guidelines for alcohol intake. 6. Document your intervention. Scenario: To address the positive depression screen, you assess their risk of suicide. 7. Document your intervention Scenario: To address the positive fall risk screen, you do a Timed Up and Go test, noting a result of 12 seconds and shuffling gait. You refer the patient to physical therapy. 8. Document your observations and intervention.
Documentation aid: comorbidities	Form for simplifying addition of common comorbidities to the problem list for inpatients	A vs. B	Scenario A: E. Z-test. Patient is being admitted for stroke. He has paroxysmal atrial fibrillation, NASH, and Stage 4 CKD. 3. Document the comorbidities using the stroke admission order set. Scenario B: A. Z-test. Patient is being admitted for hip fracture. He has altered mental status, hyponatremia, and acute kidney injury. 4. Document the comorbidities using the hip fracture admission order set.
Interruptive alert: acute opioid	Order entry alert to promote compliance with state regulations limiting prescription opioids for acute pain	A vs. B	Scenario A: H. Z-test needs a 10 d prescription for an acute opioid because of complications after a procedure. You decide to write acetaminophen/hydrocodone 5/325 1 PO Q 6 h prn x 10 d. 1. Write the prescription and document the reasons for the extended prescription. Scenario A: The patient later gets acetaminophen/oxycodone from urgent care. She calls for a refill (a second prescription) for acetaminophen/oxycodone 5/325 1 PO Q6 h prn x 10 d. 2. Write the prescription and document that it is not the initial prescription for acetaminophen/ oxycodone. Scenario B: B. Z-test needs a 10 d prescription for an acute opioid because of complications after a procedure. You decide to write acetaminophen/hydrocodone 5/325 1 PO Q6 h prn x 10 d. 3. Write the prescription and document the reasons for the extended prescription. Scenario B: The patient later gets acetaminophen/oxycodone from urgent care.

(Continued)

Table 1 (*Continued*)

Feature	Description	Testing strategy	Tasks
			She calls for a refill (a second prescription) for acetaminophen/oxycodone 5/325 1 PO Q6 h prn x 10 d. 4. Write the prescription and document that it is not the initial prescription for acetaminophen/oxycodone.
Test event #2			
Interruptive alert: drug–disease interaction checking	Alerts provider at order entry to potential medication complications based on patient problems recorded on problem list.		Scenario: For F. Z-test you wish to prescribe an oral antibiotic and your first choice is azithromycin, but doxycycline would also work. 1. Write the prescription and work through the alert. Scenario: You also recommend something for pain. Your first choice is meloxicam, but if there are issues, you'll go with over the counter acetaminophen. 2. Write the prescription and work through the alert.
Passive alert: malnutrition	Alerts provider to malnutrition diagnoses identified by dietitian		C. Z-test is a chronically ill smoker with COPD who you have been caring for on the inpatient service for several days with a COPD exacerbation. 1. What do you see?[a] Scenario: You agree with the dietitian's assessment of malnutrition. 2. Document your impression. Scenario: A. Z-test is a Type 1 diabetic hospitalized with elevated blood sugars. She has been on your service several days. 3. What do you see?[a] Scenario: You are not sure you agree with the dietitian's diagnosis of unintentional weight loss and you want to talk to the dietitian. 4. Contact the dietitian.
Clinical pathway: vascular access	Algorithm to assist providers with choosing best method of venous access, interruptive version	A (vs. B in test event 3)	Scenario: W. Z-test is in the hospital for a septic joint. The joint has been drained and now you plan 4 wk of IV antibiotics. 1. Order a PICC line and work through the alert.
Clinical pathway: apnea in stroke	Algorithm to promote screening ischemic stroke patients for obstructive sleep apnea, interruptive version	A (vs. B in test event 3)	Scenario: C. Z-test is being discharged from the hospital after an ischemic stroke. 1. Place discharge orders using the stroke discharge order set. Scenario: Scenario: D. Z-test is being discharged from the hospital after an ischemic stroke. 2. Place discharge orders using the stroke discharge order set.
Clinical pathway: neonatal syphilis	Algorithm to assist with evaluation and management of infants with possible or confirmed syphilis		Scenario: Baby Z-test is in the newborn nursery and mother tested positive for syphilis IgG and IgM. According to the reverse testing algorithm, maternal RPR was obtained and was positive. You cannot find documentation that mother was treated. No information available about the partner. Additionally, you find a normal physical exam, normal CBC and CSF, but the infant RPR compared to mother's is four-fold higher. (Repeat data elements as needed for the participant). 1. Use the pathway to determine the correct assessment and treatment. Scenario: Baby Z-test is in the newborn nursery and mother tested positive for syphilis IgG and IgM. According to the reverse testing algorithm, maternal RPR was obtained and was positive. Mother was treated adequately 8 wk prior to delivery. Partner is healthy and no concern for maternal reinfection or relapse. Additionally, you find a normal physical exam and the infant RPR equal compared to mother's is equal. CBC and LP are felt to be not indicated. (Repeat data elements

Table 1 (*Continued*)

Feature	Description	Testing strategy	Tasks
			as needed for the participant). 2. Use the pathway to determine the correct assessment and treatment.
Interruptive alert: pediatric sepsis	Alerts providers to changes in patient parameters potentially indicative of sepsis in pediatric patients		Scenario: U. Z-test is a 14-year-old with 3 d of RLQ pain and low grade fever. On exam she is lying still on the gurney; she has RLQ tenderness and rebound. She is admitted to your service for possible appendicitis. CBC, CMP, CT abdomen are ordered. One hour later you open her chart to see if admission labs are back yet (they aren't). Close the chart. 1. What do you see?[a] 2. Work through the alert.[a]
Test event #3			
Clinical pathway: vascular access	Algorithm to assist providers with choosing best method of venous access, non-interruptive version	B (vs. A in test event 2)	Scenario: S. Z-test needs a line for 3 wk of IV antibiotics. You plan to order a PICC line from the IV Therapy Placement team. 1. Add the order (to the scratchpad). 2. Find the latest eGFR. 3. Perform any other actions needed in preparation for signing the order. Scenario: W. Z-test needs a line for 3 wk of IV antibiotics. You plan to order a PICC line from the IV Therapy Placement team. 4. Add the order (to the scratchpad). 5. Find the latest eGFR. 6. Perform any other actions needed in the preparation for signing the order.
Clinical pathway: apnea in stroke	Algorithm to promote screening ischemic stroke patients for obstructive sleep apnea, non-interruptive version	B (vs. A in test event 2)	Scenario:C. Z-test is hospitalized after an ischemic stroke. You plan to discharge the patient using the stroke discharge order set. 1. Add the order (to the scratchpad). 2. Perform any other actions needed in the preparation for signing the order. Scenario: D. Z-test is hospitalized after an ischemic stroke. You plan to discharge the patient using the stroke discharge order set. 1. Add the order (to the scratchpad). 2. Perform any other actions needed in the preparation for signing the order.
Passive alert: guardianship	Alerts provider to contact guardian with changes in patient status		Scenario: C. Z-test is in the hospital after an ischemic stroke. 1. Determine if the guardian(s) has been notified about the most recent significant change in status and what that change in status was. 2. Find the phone number to use if you were making the call to the guardian(s). Scenario: You notify the guardian G. Z-test that the patient has been transferred to the floor. 3. Find the place to document the guardian notification. 4. Complete the documentation of the call with the guardian. Scenario: D. Z-test is in the hospital after an ischemic stroke. Open the chart. 5. Determine when the guardian was last notified and by whom. 6. Find the phone number to use to call the guardian.
Passive alert: 1-y mortality risk	Prompts palliative care, advance care planning, and palliative care consultations in patients with increased risk of 1-y mortality		Scenario: V. Z-test is in the hospital after an ischemic stroke. Open the chart. 1. Find options for managing or managing the patient's elevated 1-y mortality risk. 2. Order a palliative care consult. Scenario: A. Z-test is admitted to your service. 3. Find the options for managing the patient's elevated 1-y mortality risk

(*Continued*)

Table 1 (*Continued*)

Feature	Description	Testing strategy	Tasks
			4. Find the date of the last Advance Care Planning note. Scenario: You discuss Advance Care Planning and review the existing plan with the patient and family. You determine no changes need to be made. 5. Document your review.
Order: calendar icon	Alternative UI control used to postdate orders		Scenario: A. Z-test is admitted to your service. 4. Order a chest X-ray postdated 4 wk. 5. Order a complete blood count postdated 4 wk. 6. Order basic metabolic profile postdated 6 wk.

aDenotes tasks evaluated only qualitatively.

Table 2 Usability testing participants

Participant	Position	Specialty	Age	Gender	Race	Ethnicity	Years in profession	Years on this EHR
Test event 1								
1	Attending	Family Medicine	50–59	f	w	nh	33	2
2	Attending	IM/Peds	40–49	m	w	nh	11	18
3	Resident	Psychiatry	30–39	m	a	nh	2	2
4	Fellow	Palliative Care	30–39	m	b	nh	4	1
5	Resident	Family Medicine	20–29	m	w	nh	1	1
6	Physician Assistant	Family Medicine	30–39	f	w	nh		2
Test event 2								
1	Attending	Family Medicine	40–49	f	w	nh	16	17
2	Attending	Family Medicine	20–29	m	w	nh	4	0.3
3	Resident	Family Medicine	20–29	f	w	nh	1	3
4	Resident	Family Medicine	30–39	f	w	nh	2	2
5	Attending	Family Medicine	30–39	f	w	nh	5	4.5
6	Resident	Family Medicine	20–29	f	w	nh	0.25	0.25
Test event 3								
1	Resident	Family Medicine	20–29	f	w	nh	1	5
2	Resident	Emergency Medicine	20–29	m	w	nh	2.5	4.5
3	Attending	Internal medicine	30–39	m	a	nh	15	10
4	Resident	Transitional	20–29	f	w	nh	0.5	0.5
5	Attending	Pediatrics	40–49	f	w/ai	nh	16	17

Abbreviations: a, Asian; ai, American Indian; f, female; IM, internal medicine; m, male; nh, non-Hispanic; w, white.

were identified, a rate of 2.5 concerns per feature tested (range 0–9, median 2, ►**Supplementary Table S1**, available in the online version). Twenty-three (70%) of the usability concerns were judged correctable. Summative test results generally supported the impressions gathered using think aloud technique and retrospective probing (►**Supplementary Table S2**, available in the online version). Of the 13 features tested, we highlight three features in detail to illustrate a range of testing approaches, the clarity that user testing can bring to imple-

mentation decisions, and the measurable impact testing can have for end users.

Comorbidities Documentation Aid

A comobidities documentation aid was designed to simplify addition of common comorbidities to the problem list for inpatients (►**Fig. 1**). The documentation aid was included in the admission order sets. Two versions were designed, one exposing checkboxes for every possible comorbidity

Fig. 1 Comorbidities documentation aid, version B, used in production.

selection, and the other using accordion controls, initially exposing a limited number of comorbidity options. The two configurations were compared using A versus B configuration testing during the same test event. The user tasks for each configuration were as follows:

Scenario A: Edtest Z-test. Patient is being admitted for stroke. He has paroxysmal atrial fibrillation, NASH, and Stage 4 CKD.

1. Document the comorbidities using the stroke admission order set.

Scenario B: Ape Z-test. Patient is being admitted for hip fracture. He has altered mental status, hyponatremia, and acute kidney injury.

2. Document the comorbidities using the hip fracture admission order set.

Mean task time was relatively long, and completion rates and error rates reflected mixed results; however the SEQ scores were 2.5 and 2.8, among the lowest of any tasks across all test events. One option was chosen and moved to production, however, based on user feedback postimplementation, it was subsequently withdrawn after firing 2,473 times in 6 months. Based on task times from user testing, we estimate users spent a total of 50 hours in the live environment on an ineffective feature with poor usability.

Vascular Access Clinical Pathway

A clinical pathway aimed to guide users toward the most appropriate vascular access method usability tested before implementation. The pathway displayed when ordering peripherally inserted central catheters (PICC) lines. In the initial test event, the pathway opened in an interruptive modal dialog. The user scenario and task were:

Scenario: Whale Z-test is in the hospital for a septic joint. The joint has been drained and now you plan 4 weeks of IV antibiotics.

1. Order a PICC line and work through the alert.

Three usability concerns deemed uncorrectable were identified. In one instance, the application crashed while testing the feature. Another user found a particular alternative workflow which resulted in an endless loop and prevented the user from returning to the patient chart. Based on these results, the pathway was redesigned and rebuilt such that it displayed in the standard non-modal orders window (►**Fig. 2**). In the following test event, two tasks were configured for the new pathway design. We learned from the initial test that user interaction with the pathway was complex, and that we may learn more if we broke the initial test event task into smaller tasks in subsequent testing:

Sue Z-test needs a line for 3 weeks of IV antibiotics. You plan to order a PICC line from the IV Therapy Placement team.

1. Add the order (to the scratchpad).

2. Find the latest eGFR.

3. Perform any other actions needed in preparation for signing the order.

Whale Z-test needs a line for 3 weeks of IV antibiotics. You plan to order a PICC line from the IV Therapy Placement team.

4. Add the order (to the scratchpad).

5. Find the latest eGFR.

6. Perform any other actions needed in the preparation for signing the order.

Fig. 2 Second iteration of the vascular access pathway.

The redesigned pathway was viewed favorably by most participants. This assessment was backed by the summative test results. Because the redesigned pathway testing used three smaller tasks instead of the one larger task used in the initial testing, the quantitative data from these three smaller tasks were combined in order to compare the results from the initial test event. The aggregate task times in the test of the revised pathway were 140.8 seconds and 59.8 seconds, substantially lower than the initial test event task time of 287.2 seconds. Total errors dropped from 9 to 3. SEQ scores improved from 2.7 to 4.2 or greater. More usability concerns were identified with the second version of the pathway, however, these issues were mostly correctable. Testing of the second iteration of the pathway showed the pathway to be stable. The second iteration (►**Fig. 2**) was chosen for implementation and remains in use to date.

The first iteration of the vascular access clinical pathway feature was withheld from production based only on our site-specific user testing, allowing for redesign and retesting. At our institution in 2020, approximately 18 PICC lines were ordered each week. Two of the six participants testing the first version of the vascular access clinical pathway uncovered workflows leading to uncorrectable, irrecoverable errors. Withholding the venous access clinical pathway feature from production may have prevented as many as six irrecoverable errors per week that otherwise would have occurred until the feature was inevitably withdrawn.

Calendar Icon

In the original implementation of Cerner PowerChart at MU Health Care, no method of postdating orders was available from the vendor. A dropdown menu was added to all orders to accommodate the "future order" use case. Later the vendor developed a calendar icon feature which improved upon, and could replace, our original dropdown menu implementation. This calendar icon launched a modal window in which the user can specify timeframes for future orders in a variety of different ways. We user-tested the new, vendor-supplied calendar icon to assess implementation decisions and how readily users would convert to the new method of postdating orders. The tasks were as follows:

Scenario: Alligator Z-test is admitted to your service.
1. Order a chest X-ray postdated 4 weeks.
2. Order a complete blood count postdated 4 weeks.
3. Order basic metabolic profile postdated 6 weeks.

Testing revealed generally low completion rates, relatively high error rates, but good SEQ scores. The calendar icon and related modal window are not modifiable by sites, but the position of the icon within the order fields is configurable by site. The one usability concern found related precisely to this configuration. At certain lower screen resolutions, the calendar icon could not be seen in the order entry fields without scrolling. Participants mistakenly used an existing date field with a different purpose but similar name to enter the intended future order date (►**Fig. 3**). The solution for this error is reconfiguration of the order fields and labels (►**Fig. 4**).

Discussion

Descriptions of discount usability in EHR system evaluations generally focus on either heuristic evaluations,[18–20] a specific feature under test,[21,22] or both.[19,23] We are aware of no descriptions in the literature of the broad application of site-specific, discount user testing of a health care system's local configuration, despite the potential for such configurations to substantially impact usability. This case report highlights the value of user testing using principles of discount usability at MU Health Care. The existing UCD process already leveraged the discount usability methods of heuristic evaluations and iterative design.[12,13] We found that a user test program, incorporating simplified user testing with a limited number of participants and use of the think aloud technique, supplements those UCD methods.

We recorded standard usability metrics such as completion rates, task times, and errors. These are not elements of discount usability.[12] However, the most important findings from our user testing were derived from the think aloud technique and test moderator observations. The advanced metrics often supported the qualitative findings in our testing, but at times those findings were mixed, and in general they did not drive important design or implementation decisions. Deriving the quantitative results was particularly time consuming because it required a second review of
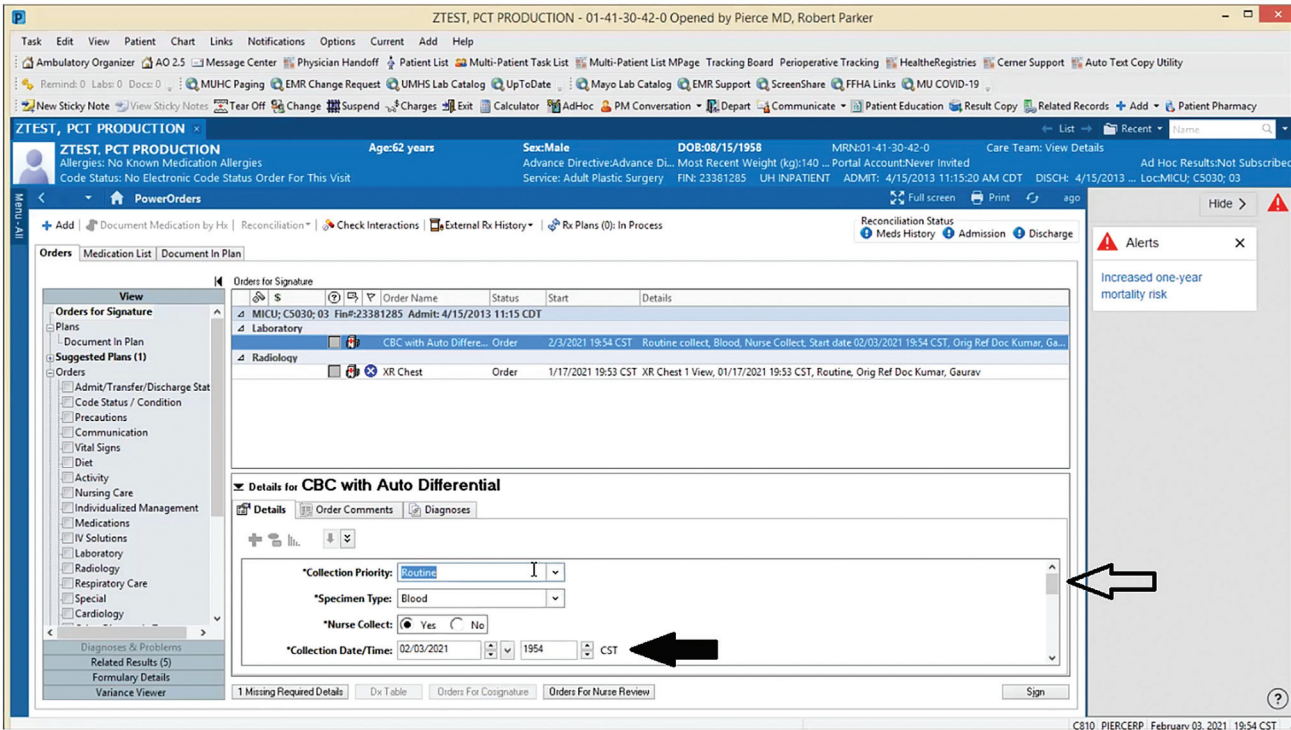
Fig. 3 Screen recording of user testing of the calendar icon. At lower screen resolutions, the calendar icon is not accessible without scrolling (*open arrow*). The collection date/time field (*solid arrow*), a related field with a similar name but with a different purpose, is higher in the order of the fields. Used with participant permission (Note: Fictional patient data).
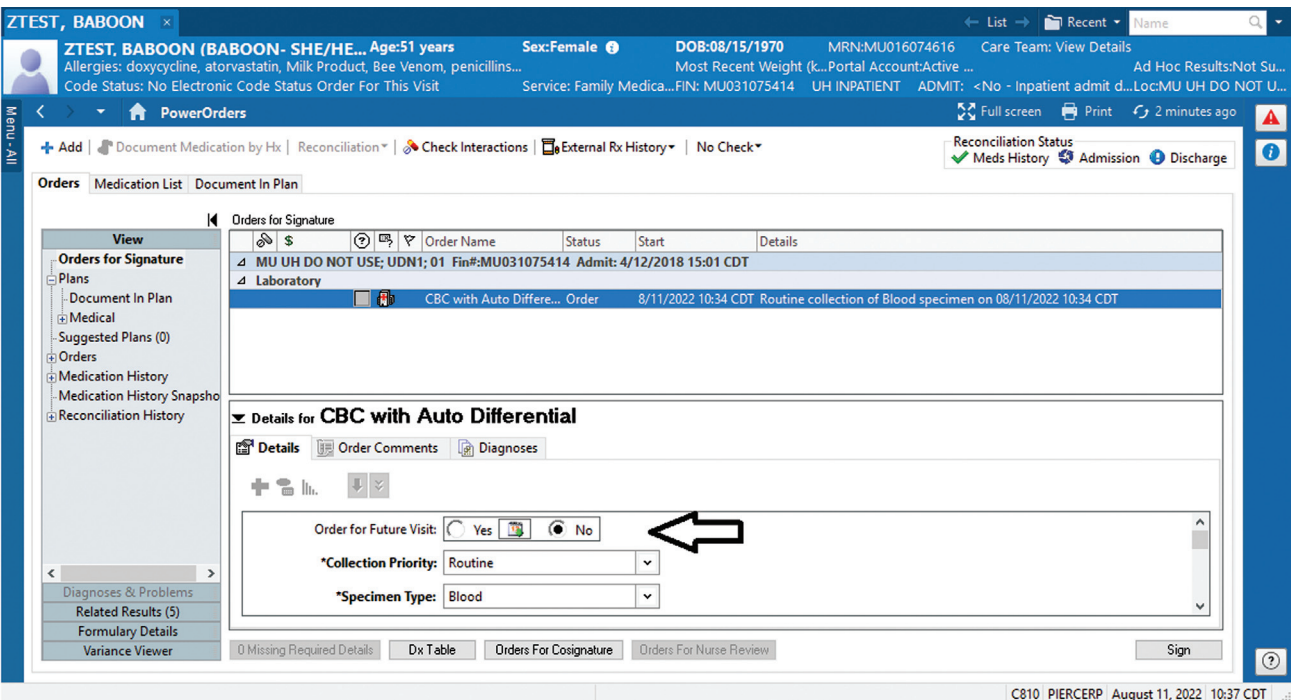
Fig. 4 Calendar icon repositioned based on usability testing (*open arrow*) (Note: Fictional patient data).

the recordings. We found some value in the simple and easily obtained SEQ score, which in the case of comorbidities documentation aid appears to have portended the withdrawal of that feature from production. Inclusion of other advanced usability metrics such as task time and error rates, which are not elements of discount usability, did not greatly augment our evaluations. In site-specific user testing programs of EHR configurations, advanced usability metrics may not provide enough value and may limit the scalability of such programs.

The usability methods described in this report align with a recent systematic review of health-related usability evaluations.[24] Among the methods recommended by experts and employed in our program include remote user testing (combined with think aloud or interview), heuristic evaluation, and A/B testing. In an example not highlighted in this report, A/B testing gave a clear indication of a preferred option for an acute opioid clinical decision support alert in the first test event. Site-specific user testing of local configurations need not be limited to discount usability methods. Other methods, such as A/B testing, may have value for health care systems.

The features highlighted in this report were chosen to highlight the spectrum of usability concerns found with a site-specific user testing program. The original vascular access pathway design did not account for the alternative workflows found in the user test event, and complete redesign of this feature likely prevented a large number of unfortunate user experiences. On the other hand, the calendar icon provided by the vendor could not be redesigned. We had configuration control only over where in the order the icon appeared. Changes to these fields are relatively simple, but the need for these changes was only made apparent through our user testing program.

For the first test event we used Morae, an application specifically designed for user testing. The software is complex, takes time to learn and configure, and is difficult to use with remote testing. We changed to an online meeting platform for the second and third test events. The health care organization's rapid adoption of this platform as a result of the COVID-19 telehealth expansion meant both moderators and participants were very familiar with it. The move to a lighter, simpler platform meant less quantitative data but also simpler configurations, less prep time for each testing session, and more rapid test cycling. The forced change from Morae to Zoom served to improve our program by shifting the focus from quantitative to qualitative findings, a key component of discount usability.

Health care organizations may be challenged to provide enough resources to test their site-specific implementations.[25] Recruiting test participants for the first test event was difficult. A monetary incentive helped a great deal, and ongoing support for participant incentives may be needed to ensure consistent participation. Configuring test patients, preparing test materials, analyzing results, and preparing reports take considerable time, and support from health care organization leadership is essential. Most importantly, health care organizations may not have staff with training or experience in usability evaluations. The proper design, conduct, and interpretation of usability tests is aided by familiarity with usability evaluation methods and concepts such as learnability, efficiency, effectiveness, and satisfaction. Even the seemingly simple process of writing good tasks can be challenging.[26,27] And experience does matter: higher quality usability methodology does detect more usability problems.[13] Yet Jakob Nielsen himself notes that "…even people who are not dedicated usability specialists…can still conduct user studies" and that "bad user testing beats no user testing."[12] We acknowledge that our user testing program is imperfect, but our approach, using elements of discount usability, has served to improve the user experience and may provide a template for other health systems to implement site-specific user testing programs of their own.

Further research may clarify the best practices and additional ways to make site-specific user testing scalable so that users benefit from testing of all aspects of an implementation. We did not usability test every implemented feature. Work which identifies which types of features introduce the most serious usability concerns would help sites focus testing efforts to maximize value. We did not collect System Usability Scale (SUS) scores.[28] The SUS is a commonly used and well known usability metric, but the ten-item SUS was too time consuming to use after each feature. A recent review cites the SUS as an example of a questionnaire "not recommended" as an EHR usability evaluation method.[24] Whether and in what contexts the SUS adds value to site-specific user testing remains to be determined. Finally, testing outside of formal "test events" may help an organization scale up a testing program by simplifying recruitment. Future work should define the value of testing in other settings such as simulation labs and EHR help centers.

This study has important limitations. First, we made two important changes after the first test event. Adding a financial incentive likely altered recruitment patterns and user testing using incentives. The switch to an alternative recording software program could have altered quantitative usability test results. Because of these changes, aggregate results should be interpreted with caution. Second, this report was limited to user testing with providers. Nurses may have different usability concerns and the yield from a discount usability approach may differ from that described in this report. Finally, this is a single center study, limiting its generalizability. Other health care systems with different resources and experience may see different results from the addition of user testing to their UCD methods.

Conclusion

Our work suggests that a site-specific user testing program can prevent the implementation of features with poor usability and aid in the identification of correctable usability problems. Given limitations in time, training, and other resources, a discount usability program like the one described in this report likely represents what is currently "in reach" for user testing of site-specific implementations for many health care organizations. We recommend that systems without any UCD processes consider the simplest discount usability methods such as heuristic review, paper prototyping, and iterative design; systems already using such methods should consider addition of a user testing component, emphasizing think aloud technique, limited numbers of participants, and limited use of advanced quantitative metrics. Additional work is needed to further characterize the role and benefits of a testing program and describe the optimal use of limited system resources for the greatest improvement in user experience.

Clinical Relevance Statement

Site-specific user testing identifies superior feature designs and a number of correctable usability concerns representing all types of heuristic violations. Site-specific user testing programs may supplement existing UCD methods and improve the EHR experience for the end user.

Multiple Choice Questions

1. Which of the following is essential for performing discount site-specific user testing?
   a. Team member(s) with extensive training in usability.
   b. Assistance from the EHR vendor.
   c. Team member(s) with enthusiasm for user testing.
   d. Software specifically designed for usability testing.

   Correct Answer: The correct answer is option c. Given limitations in time, training, and other resources, a discount usability program like the one described in this report likely represents what is currently "in reach" for user testing of site-specific implementations for many health care organizations.

2. Benefits to performing site-specific user testing may include which of the following?
   a. Improving end-user EHR experience.
   b. Increasing task completion times.
   c. Widening the her usability gap.
   d. Implementing features with heuristic violations.

   Correct Answer: The correct answer is option a. Improving end-user EHR experience. Specific user testing programs may supplement existing UCD methods and improve the usability of the EHR for the end user.

Protection of Human and Animal Subjects
The project details were reviewed by the University of Missouri Institutional Review Board who determined the project to be a quality improvement activity and not human subject research and did not require additional review.

Conflict of Interest
The authors declare that they have no conflict of interest or competing interests in the project. MU Health Care was responsible for all aspects of testing and did not receive support, training, or other assistance from the EHR vendor. B.R. is employed by Cerner Corporation, the vendor, but her sole responsibilities are for the MU Health Care site-specific implementation.

References

1 Office of the National Coordinator for Health Information Technology Health IT Quick Stats. U.S. Department of Health and Human Services. Accessed August 05, 2021 at: https://dashboard.healthit.gov/quickstats/quickstats.php
2 International Standards Organization. Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. International Organization for Standardization 2019
3 Melnick ER, Dyrbye LN, Sinsky CA, et al. The association between perceived electronic health record usability and professional burnout among US physicians. Mayo Clin Proc 2020;95(03):476–487
4 Committee on Patient Safety and Health Information Technology Institute of Medicine. Health IT and Patient Safety: Building Safer Systems for Better Care. National Academies Press (US) 2011
5 Howe JL, Adams KT, Hettinger AZ, Ratwani RM. Electronic health record usability issues and potential contribution to patient harm. JAMA 2018;319(12):1276–1278
6 US Department of Health and Human Services. Health Information Technology: Standards, Implementation Specifications, and Certification Criteria for Electronic Health Record Technology, 2014 Edition; Revisions to the Permanent Certification Program for Health Information Technology. In: Federal Register, ed.; 2012
7 US Department of Health and Human Services. 2015 Edition Health Information Technology (Health IT) Certification Criteria, 2015 Edition Base Electronic Health Record (EHR) Definition, and ONC Health IT Certification Program Modifications. In: Federal Register, ed.; 2015
8 Tutty MA, Carlasare LE, Lloyd S, Sinsky CA. The complex case of EHRs: examining the factors impacting the EHR user experience. J Am Med Inform Assoc 2019;26(07):673–677
9 Ratwani RM, Savage E, Will A, et al. A usability and safety analysis of electronic health records: a multi-center study. J Am Med Inform Assoc 2018;25(09):1197–1201
10 Meigs SL, Solomon M. Electronic health record use a bitter pill for many physicians. Perspect Health Inf Manag 2016;13(Winter):1d
11 Ratwani RM, Sinsky CA, Melnick ER. Closing the Electronic Health Record Usability Gap. Bill of Health blog. June 26, 2020, 2020. Accessed February 8, 2021 at: https://blog.petrieflom.law.harvard.edu/2020/06/26/closing-the-electronic-health-record-usability-gap/
12 Nielsen J. Discount Usability: 20 Years. Nielsen Norman Group; 2022
13 Nielsen J. Usability Engineering at a Discount. Elsevier Science Publishers; 1989
14 Customized Common Industry Format Template for Electronic Health Record Usability Testing (NISTIR 7742) National Institute for Standards and Technology. 2010:1–37
15 TechSmith. TechSmith Support Policy. TechSmith 2021. Accessed August 05, 2021, at: https://support.techsmith.com/hc/en-us/articles/203732728
16 Sauro J, Dumas JS. Comparison of three one-question, post-task usability questionnaires. Paper presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2009 Boston, MA. Doi: 10.1145/1518701.1518946
17 Nielsen J. 10 Usability Heuristics for User Interface Design. Nielsen Norman Group Accessed May 17, 2020 at: https://www.nngroup.com/articles/ten-usability-heuristics/
18 Corrao NJ, Robinson AG, Swiernik MA, Naeim A. Importance of testing for usability when selecting and implementing an electronic health or medical record system. J Oncol Pract 2010;6(03):120–124
19 Pertiwi AAP, Fraczkowski D, Stogis SL, Lopez KD. Using heuristic evaluation to improve sepsis alert usability. Crit Care Nurs Clin North Am 2018;30(02):297–309
20 Tarrell A, Grabenbauer L, McClay J, Windle J, Fruhling AL. Toward improved heuristic evaluation of EHRs. Health Syst (Basingstoke) 2015;4(02):138–150
21 Beaudoin DE, Rocha RA, Tse T. Enhancing access to patient education information: a pilot usability study. AMIA Annu Symp Proc 2005;2005:892–892
22 Schaarup C, Hejlesen OK. Heuristic evaluation and thinking aloud test of a digitized questionnaire for diabetes outpatient clinics. Stud Health Technol Inform 2014;205:920–924
23 Khelifi M, Tarczy-Hornoch P, Devine EB, Pratt W. Design recommendations for pharmacogenomics clinical decision support systems. AMIA Jt Summits Transl Sci Proc 2017;2017:237–246

24 Sinabell I, Ammenwerth E. Agile, easily applicable, and useful ehealth usability evaluations: systematic review and expert-validation. Appl Clin Inform 2022;13(01):67–79

25 Hettinger AZ, Melnick ER, Ratwani RM. Advancing electronic health record vendor usability maturity: progress and next steps. J Am Med Inform Assoc 2021;28(05):1029–1031

26 Schade A. Write better qualitative usability tasks: top 10 mistakes to avoid. Nielsen Norman Group Accessed February 21, 2021. 2021 at: https://www.nngroup.com/articles/better-usability-tasks/

27 Russ AL, Saleem JJ. Ten factors to consider when developing usability scenarios and tasks for health information technology. J Biomed Inform 2018;78:123–133

28 Brooke J. SUS: A "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, eds. Usability Evaluation in Industry. Taylor and Francis; 1996