

# A Pilot Study on Usability Analysis of Emergency Department Information System by Nurses

M.S. Kim<sup>1,2</sup>; J.S. Shapiro<sup>2,3</sup>; N. Genes<sup>3</sup>; M.V.Aguilar<sup>4</sup>; D. Mohrer<sup>2</sup>; K. Baumlin<sup>3</sup>; J.L. Belden<sup>5</sup>

<sup>1</sup>Department of Health Management and Informatics, University of Missouri School of Medicine, Columbia, MO, USA; <sup>2</sup>Center for Biomedical Informatics, Mount Sinai School of Medicine, New York, NY; <sup>3</sup>Emergency Department, Mount Sinai School of Medicine, New York, NY; <sup>4</sup>Information Technology, Mount Sinai Medical Center, New York, NY; <sup>5</sup>Family and Community Medicine, University of Missouri School of Medicine, Columbia, MO, USA

## Keywords

Electronic health records, emergency department information system, EDIS, usability engineering

## Summary

**Objectives:** Employing new health information technologies while concurrently providing quality patient care and reducing risk is a major challenge in all health care sectors. In this study, we investigated the usability gaps in the Emergency Department Information System (EDIS) as ten nurses differentiated by two experience levels, namely six expert nurses and four novice nurses, completed two lists of nine scenario-based tasks.

**Methods:** Standard usability tests using video analysis, including four sets of performance measures, a task completion survey, the system usability scale (SUS), and sub-task analysis were conducted in order to analyze usability gaps between the two nurse groups.

**Results:** A varying degree of usability gaps were observed between the expert and novice nurse groups, as novice nurses completed the tasks both less efficiently, and expressed less satisfaction with the EDIS. The most interesting finding in this study was the result of 'percent task success rate,' the clearest performance measure, with no substantial difference observed between the two nurse groups. Geometric mean values between expert and novice nurse groups for this measure were 60% vs. 62% in scenario 1 and 66% vs. 55% in scenario 2 respectively, while there were some marginal to substantial gaps observed in other performance measures. In addition to performance measures and the SUS, sub-task analysis highlighted navigation pattern differences between users, regardless of experience level.

**Conclusion:** This study will serve as a baseline study for a future comparative usability evaluation of EDIS in other institutions with similar clinical settings.

## Correspondence to:

Min Soon Kim, PhD  
Department of Health Management and Informatics  
University of Missouri School of Medicine  
CE728 Clinical Support & Education, DC006.00  
Columbia, MO 65212  
Email: kimms@health.missouri.edu

*Appl Clin Inf* 2012; 3: 135–153

doi:10.4338/ACI-2011-11-RA-0065

received: November 1, 2011

accepted: March 4, 2012

published: March 28, 2012

Citation: Kim MS, Shapiro JS, Genes N, Aguilar MV, Mohrer D, Baumlin K, Belden JL. A pilot study on usability analysis of emergency department information system by nurses. *Appl Clin Inf* 2012; 3: 135–153  
<http://dx.doi.org/10.4338/ACI-2011-11-RA-0065>

## 1. Background

### 1.1 Scientific background

The Health Information Technology for Economic and Clinical Health (HITECH) Act and its incentives for meaningful use will likely drive widespread Electronic Health Record (EHR) adoption across the US. However, among the barriers to EHR adoption is poor usability [1, 2]. Usability is the overarching concept of the effectiveness, efficiency and satisfaction that users of specific product can achieve, to a certain defined level, with an explicit set of tasks in a specific environment [3]. Prior investigations have demonstrated that a lack of usability consideration in the design of clinical data management systems creates potential human-computer interaction issues, including increased workflow complexity that will result in loss of productivity in clinical practice and research, and decreased quality of patient care [4, 5]. For example, the Koppel study reported 22 types of medication errors facilitated by a widely used hospital computerized provider order entry system (CPOE), a cardinal functionality of EHRs, the result of which was contradictory to the prior belief that CPOE reduce medication errors [6]. It is therefore necessary to create usability evaluation standards for newly implemented EHR systems, in order to avoid certain human-computer interaction issues and improve patient care in many clinical settings [7–9].

Usability issues have also been confirmed through user satisfaction survey results by EHR end users. According to one recent user satisfaction survey, over 30% of more than 3,700 EHR user respondents would not recommend their EHR [10]. Similarly, “The 2009 EHR User Satisfaction Survey,” published in *Family Practice Management* noted that nearly 50% of 2,012 family physician respondents were not satisfied with many of the best-known EHR systems [11].

### 1.2 Rationale for the study

Effective EHR applications should not only satisfy functional capability requirements but also employ user interfaces designed to simplify the user experience, minimize the learning curve associated with human-computer interactions, and consequently streamline the clinical workflow process. In order to avoid human-computer interaction problems and improve patient care in various clinical settings, it is necessary to create usability evaluation standards for selection and implementation of EHR systems [9, 12, 13].

Due to the immediacy of data needs, the often chaotic nature of the environment, and variable volume and acuity, emergency departments (EDs) are among the most difficult areas to manage within a hospital. By integrating an Emergency Department Information System (EDIS) that can automate certain workflow processes, it is possible to minimize certain risks caused by human error, eliminate supply chain lags, and effectively expedite patient turnover while reducing costs and maintaining high quality patient care [14]. However, ED adoption of comprehensive information systems has been slow [15, 16] despite perceived advantages in the efficiency and safety brought by health information technology (IT) in other areas of the hospital [17, 18]. Nursing documentation is an integrated part of the clinical documentation and the CPOE components in an EDIS. It allows nurses to document a comprehensive clinical note and the completion of orders through the CPOE system. Accurate emergency nursing documentation is essential for continuity of care, patient safety and for hospital ED reimbursement by determining the ED level of care.

### 1.3 Study context

#### 1.3.1 Organizational setting

The Mount Sinai Medical Center is an urban, tertiary care academic medical center in New York City. The Department of Emergency Medicine has 35 full-time clinical faculty and 88 nurses who staff the clinical department. The department had 44 licensed beds and had 101,229 ED visits in 2009.

### 1.3.2 System details and system in use

The EDIS was implemented by the ED at Mount Sinai in 2003–2004. The EDIS includes physician and nursing documentation, computerized provider order entry (CPOE), results retrieval, a print-on-demand electronic prescribing solution, various modules of clinical decision support, and, through the creation of 14 electronic interfaces, comprehensive integration with hospital systems including registration, laboratory, and hospital electronic data repository. This browser based software combines triage, patient tracking, physician and nursing documentation, risk management, charge management, integrated voice recognition, prescription writing and other unique features. A multi-phase clinical workflow and process redesign effort was undertaken before EDIS implementation, with the collaboration of physician and nursing groups, to optimize the integration of the EDIS. With the support of IT leadership, the EDIS has been continued to undergo additional customization over time to reflect the specific needs of both clinician groups. As a result, the implementation of the EDIS at Mount Sinai has shown substantial improvement in the ED's operational and financial efficiency [19, 20]. However, there has been no formal usability testing of the EDIS, which could help to further improve the system and increase its beneficial effects on clinical workflow. To our knowledge, no comprehensive usability test of any EDIS by on-site clinicians has previously been described.

## 2. Objective

In this study, a standard usability test of performance metrics, a task completion survey and a system usability scale were employed. Quantitative and qualitative usability data were collected from ten nurses, comprising two groups based on level of clinical and EDIS experience in the ED, with each nurse performing two scenario-based sets of nine clinical tasks in a laboratory setting. We aimed to investigate the usability gaps in the EDIS between the two nurse groups using a series of usability evaluation measures. In doing so we created a comprehensive set of clinical functionality tasks for usability evaluation in the ED that spanned from the entry of the patient through door to final disposition decision (i.e. discharge or hospital admission).

## 3. Methods

### 3.1 Study design

Standard usability testing was conducted using a usability evaluation software package (Morae, TechSmith, Okemos, MI), as ten emergency nurses completed two sets of artificial scenario-based tasks in a laboratory setting. Quantitative and qualitative analyses, including four sets of performance measures, a task completion survey, system usability scale (SUS) measurement, and sub-task analysis were applied to examine the gaps between two nurse groups. In this study, the EDIS user interface was displayed on a standard 22 inch LCD monitor with 1024 by 768 pixel resolution using Windows XP with the Internet Explorer 8 browser. All experiments were run in the research laboratory of the Center for Biomedical Informatics at Mount Sinai School of Medicine. The usability evaluation trials were conducted in hour-long sessions. The evaluation room consisted of one participant nurse, a facilitator (MSK), and a data logger (DM) throughout the study to maintain consistency. At the beginning of the evaluation session, the participant signed an informed consent form, acknowledging that: participation is voluntary, and that volunteers have the right to stop participating at any time. The participant was then instructed to read the printed task description and was encouraged to think aloud as he/she completed each task (► Table 1 and ► Table 2). The facilitator sat next to the participant providing oversight of the evaluation session and guiding the participant throughout the session. In the usability evaluation, both the facilitator and data logger observed and documented user behavior, user comments, and system actions during the session and discussed these issues with the participant during the debriefing session for clarification purposes. This study was reviewed and approved by the Mount Sinai Institutional Review Board.

### 3.2 Participants

There has been much debate about how many participants are needed in a usability test to reliably identify usability issues [21]. In this pilot study, we initially intended to identify at least five participants for each nurse groups because, based on a review of the literature, in most usability studies this is a sufficient number to uncover the most important usability issues [22, 23]. In addition, the discovery rate of usability issues based on an estimate of return on investment is sufficient with use of 5 participants [24–26]. After careful discussion with physicians (JSS, NG) and nurse (MVA) on our team, who are themselves clinical and EHR experts, we decided to use two criteria, clinical training level and experience with the EHR, to differentiate participants into two groups. There were 65 nurses that met the expert definition available when the study was conducted. Taking into account various staffing and scheduling issues, we conducted many of the sessions during morning hours when the ED is less crowded. Ten full-time staff nurses were approached based on availability on the clinical schedule and with the approval of the clinical coordinator, and none declined to participate in the study. Ages of the participants ranged from 25 to 58. One of 10 participants was male. Because of the small sample size for this type of usability study, we did not attempt to control for age or gender. Based on the criteria, newly hired nurses on the job for less than two months were categorized as novice users. Nurses who were working in the ED using the same EHR for more than 2 years were categorized as expert users. While it was relatively easy to identify the defined experienced nurses, it was difficult to find novice nurses since their availability depended on the ED hiring plan. The expert group's user interaction with the application was used as the baseline. Usability test results from the novice group were compared with those of experts for gap analysis. Before the study, participants were assigned pseudonyms indicating clinical roles and experience levels. Information obtained was recorded in a de-identified manner and no attempts were made to re-identify participants throughout the study.

### 3.3 Test scenarios

Both physician and nurse groups use the EDIS application; however, they have different clinical responsibilities and therefore use different templates provided by the application. Thus, only clinical tasks commonly performed by primary nurses were selected. Similarly, there are differences between expert and novice nurse groups. For example, only expert nurses perform triage, and were therefore allowed to complete 'triage nurse tasks', while both expert and novice nurses complete 'primary nurse tasks'. In this study, two sets of classic cases were presented: appendicitis, which typically leads to hospital admission (► Table 1) and a left hand laceration, which usually leads to discharge once treated with a minor procedure (► Table 2). Accordingly, two lists of nine tasks, commonly performed by both expert and novice primary nurses were generated to have the subjects run through a fairly comprehensive set of clinical functionalities including ordering and documentation of medications and procedures. The tasks were fairly well constrained by an obvious diagnosis with a clear treatment pathway that both nurse groups could follow in order to avoid excessive clinical cognitive challenges or ambiguity. The tasks were coded by task sequences and grouped by functions. The tasks were:

- **Task 1** Log in
- **Task 2** Enter allergy/current meds
- **Task 3** Complete PMH
- **Task 4** Nursing Assessment
- **Task 5** Document procedure with option to enter vital signs
- **Task 6** Order diagnostic tests
- **Task 7** Document results of tests
- **Task 8** Record follow up/patient's response to medications given
- **Task 9** Complete discharge/admission pathway

### 3.4 Usability analysis

The fundamental purpose of a usability evaluation is to quantitatively and qualitatively measure the effectiveness and efficiency of the clinical application as well as the satisfaction of its users as they

perform a series of tasks. Accurate and complete capture of user behavior is critical for retrospective analysis and rapid usability feedback. The Morae software package records desktop activity on the user's computer with a video camera to record the user's facial expression, and captures all system events, including mouse clicks, Web page changes, onscreen text, error types and counts, participant comments and more. While recording the video session, the data logger (DM) coded the important instances (questions, comments, difficulties, helps, errors, etc.) with markers placed in Morae for later video analysis. Basic video analysis was employed, which required approximately 3 hours of analysis for every hour of recording. Video sessions were segmented to individual tasks and sub-tasks for an initial analysis. In this first phase of analysis, the previously marked instances during the recording session were reviewed and any missing instances were added. When there were certain instances of interest for further scrutiny, a more careful granular analysis was used in subsequent video analysis to reveal usability problems.

## 3.5 Outcome measures

### 3.5.1 Performance measures

There is no single measure that completely accounts for user performance. Thus, it is of importance to apply multiple measures working in concert to elicit different usability perspectives. Performance metrics used in the study included:

1. Percent task success rate – measures the percentage of a given task that participants successfully complete without critical errors. These errors result in an incorrect or incomplete outcome
2. Time-on-task – measures the time taken to complete a given task, measured from the time the evaluator clicks “begin task” to when he or she clicks “end task”
3. Mouse clicks – measures the number of clicks to complete a task and
4. Mouse movement – measures the length of the navigation path to complete a task.

Mouse click is a raw count of clicks to complete a given task, while mouse movement is a pixel accumulation to measure how far the participant travels to complete a task. For both measures, lower values usually indicate higher performance. Alternatively, higher values usually indicate lower performance because this may represent difficulties, stress and concerns the user may have with the application. Consequently, it is important to minimize the mouse clicks and movements to improve performance.

### 3.5.2 Task completion survey

Post-task ratings of difficulty in a usability test have the potential to provide diagnostic information and be an additional measure of user satisfaction [27]. Participants were asked to rate difficulty of task completion immediately after completing each task using a five-point ordinal scale with 1 indicating very easy and 5 indicating very difficult.

### 3.5.3 System Usability Scale

In addition to performance measures and task completion survey, we applied a system usability scale (SUS) [28], a simple, ten-item survey that provides a comprehensive assessment of subjective usability. This was completed by each participant at the end of each artificial scenario, but before any debriefing took place. Introduced in 1984, SUS has been one of the most widely accepted subjective ratings tools. This scale has been thoroughly evaluated for its reliability and validity [29–31]. Previous studies employing the SUS in healthcare include: internet based diabetes management software [32], clinical decision support tool for osteoporosis disease management [33], semantic health information publication system in Finland [34]. In addition, the Certification Commission for Health Information Technology (CCHIT), an Office of the National Coordinator – Authorized Testing and Certification Body (ONC-ATCB) for EHR technologies adopted SUS as one of the component usability measures for ambulatory EHR systems [35]. National Institute of Standards and technology (NIST) has also confirmed the use of SUS to measure the usability of EHR systems in the EHR evaluation guideline [36]. In the study, participants were asked to record their immediate response to each item, rather than thinking about items for a long time. SUS yields a single number

representing a composite measure of the overall usability of the system being studied. This yields a score from 0 to 100, with 100 representing a perfect score.

### 3.5.4 Sub-task analysis

Usability analysis of EHR systems must include the sub-tasks that make up the actual work to understand how users interact with the information. In a complex EHR system, different nurses may complete the same task in unique ways. Individual tasks were segmented into sub-tasks, analyzed and compared across the participants and scenarios in order to identify types of errors and subtle workflow and navigation pattern variability.

## 4. Results

### 4.1 Usability analysis

Overall, all ten participating nurses completed the evaluation sessions without technical difficulties and provided suggestions for EDIS improvements throughout the sessions. They had little trouble in understanding usability tasks. Some usability gaps existed between expert and novice nurses, as novice nurses completed the tasks both less efficiently and less effectively, and expressed less satisfaction with the EDIS system.

#### 4.1.1 Performance measure: percent task success rate

Geometric mean was reported throughout the analysis to prevent bias from limited sample size where mean or median value are less accurate representation of data [37, 38]. Geometric mean values of percent task success rates of nine tasks test participants were compared between two nurse groups across two scenarios (► Figure 1, top). Overall, no substantial difference was observed in completing tasks between two nurse groups as geometric mean values were compared for scenario 1 (60%, expert group vs. 62%, novice group) and scenario 2 (66%, expert group vs. 55%, novice group). Both groups achieved very low success rate in tasks 5 (Document procedure with option to enter vital signs) and 8 (Record follow up/patient's response to medications given). For scenario 1, out of nine tasks the expert nurse group achieved higher success rate in five tasks: task 2, 3, 4, 5, and 9, same success rate in one task: task 1, and lower success rate in three tasks: task 6, 7, and 8. Most experts failed in task 8 (Record follow up/patient's response to medications given) and most novices failed in task 5 (Document procedure with option to enter vital signs). On the other hand, for scenario 2, the expert nurse group achieved higher success rate in only three tasks: task 2, 3, and 7; same rate in two tasks: task 1, 6, and 8, and lower success rate in three tasks: tasks 4, 5, and 9. Task 8 was not fulfilled by novices nor experts, while only novices failed in task 3. Debriefing sessions with the participants and the sub-task analysis found that most experts claimed they knew how to document properly but in reality they often either documented inappropriately using free text in the wrong part of the note and, or gave up because of an unfamiliar layout caused by a unique macro, a customizable EHR documentation template used to create a chart of pre-formatted text. In other cases, two expert nurses were confused between "medication documentation" in the follow up section and "medication administration service" because the terms were non-intuitive. The nurses were frustrated and gave up completing this task.

#### 4.1.2 Performance measure: time on task (TOT)

Geometric mean values of time-on-task (TOT), the time taken to complete a given task, were compared between the two nurse groups (► Figure 1, bottom). Shorter TOT indicates better performance. Unlike percent task success rate results, substantial difference was observed in times on task between two nurse groups for scenarios 1 (85s, expert group vs. 163s, novice group) and scenario 2 (84s, expert group vs. 133s, novice group), respectively, with novice nurses spending 1.5–2 times longer than expert nurses. For scenario 1, out of nine tasks the expert nurse group achieved lower TOT in seven tasks: tasks 2, 3, 4, 6, 7, 8 and 9, and higher TOT in two tasks: task 1 and 5. Similarly, for scenario 2, expert nurse group achieved lower TOT in eight tasks: tasks 1, 2, 4, 5, 6, 7, 8 and 9 and higher TOT in only in the task 3.

### 4.1.3 Performance measure: mouse clicks

Geometric mean values of mouse clicks, measure of the number of clicks to complete a given task were compared between the two nurse groups (►Figure 2, top). Lower values indicate better performance. Overall, the expert nurse groups completed the tasks with slightly fewer mouse clicks across the tasks for both scenarios. Both groups spent most of the clicks on task 4: Nursing Assessment and task 6: Order diagnostic tests for both scenarios. For scenario 1, out of nine tasks the expert nurse group achieved fewer mouse clicks in six tasks: tasks 3, 4, 6, 7, 8 and 9, and more clicks in three tasks: task 1, 2, and 5. For scenario 2, the expert nurse group required fewer mouse clicks in only four tasks: tasks 5, 6, 7, and 9 and more clicks in 2, 3, 4, and 8, and same clicks in task 1.

### 4.1.4 Performance measure: mouse movements

Geometric mean values of mouse movement, length of the navigation path to complete a given task were compared between two nurse groups (►Figure 2, bottom). Lower pixel values usually indicate higher performance. Overall, the expert nurse group showed slightly shorter mouse movements across the tasks for both scenarios. Both groups spent most of the mouse movements on task 6: Order diagnostic tests for both scenarios, which is consistent with the results of mouse clicks. For scenario 1, out of nine tasks the expert nurse group showed shorter mouse movements in eight tasks: tasks 2, 3, 4, 6, 7, 8, and 9, and longer mouse movements in task 5. Similarly, for scenario 2, the expert nurse group moved less in five tasks: tasks 1, 2, 4, 6, and 7 and more in tasks 3, 5, 8 and 9.

### 4.1.5 Task completion survey

Geometric mean values of the task completion survey on how difficult it was to complete a given task were compared between two nurse groups (►Figure 3). Values near 1 indicate 'very easy' and 5 indicate 'very difficult'. Expert nurse group showed no values above 3 (neutral) across the tasks for both scenarios, while novice groups expressed difficulty in completing tasks 6, 9 in scenario 1 and task 8 in scenario 2, respectively.

### 4.1.6 Correlation among performance measures and task completion survey

There are several studies showing mixed results regarding the effectiveness of subjective rating scales with performance measures [39,40]. Thus, it is of importance to see if the quantitative measures correlate with subjective measures to understand the effectiveness of individual metrics and if there is a systematic pattern to support or refute.

►Table 3 reports *Pearson correlation coefficient* between performance measures and subjective task completion survey ratings, computed using all scenario based tasks ( $N = 108$ , expert group and  $N = 63$ , novice group) on both nurse groups separately. Acronyms were used for succinct representation of the measures: MC (Mouse Clicks), MM (Mouse Movements), TOT (Time on Task), TCS (Task Completion Survey), PTSR (Percent Task Success Rate). Overall, weak or marginal correlation was observed between the task completion survey ratings and other performance measures for both nurse groups.

Time on task was the common highest, but still had marginal correlation for both groups ( $r = 0.40$ ,  $p = 0.00$ , expert group and  $r = 0.32$ ,  $p = 0.01$ , novice group) against the task completion survey rating. Mixed results were observed for the expert nurse group: strong positive correlation was observed between TOT vs. MM ( $r = 0.83$ ,  $p = 0.00$ ), while weak to moderate correlation was observed between TOT vs. MC ( $r = 0.30$ ,  $p = 0.00$ ), and TCS vs. TOT ( $r = 0.40$ ,  $p = 0.00$ ).

Moderate to strong correlations were observed for novice nurse group: MM vs. MC ( $r = 0.64$ ,  $p = 0.00$ ), TOT vs. MC ( $r = 0.68$ ,  $p = 0.00$ ) and TOT vs. MM ( $r = 0.74$ ,  $p = 0.00$ ), while moderate correlation was observed between TCS vs. TOT ( $r = 0.32$ ,  $p = 0.01$ ). Another observation was a very weak correlation of percent task success rate compared with all other measures, with the weakest being task completion survey ratings, which may tend to confirm the prior research that a participant may fail and still rate the task easy [41].

### 4.1.7 System Usability Scale

The SUS demonstrated that novice nurses rated the system usability at 55 (marginal) and 43 (unacceptable) and experts rated it at 75 and 81 (excellent) across the scenarios [31]. This result may in-

dicating that expert users who already achieved a certain level of proficiency using the EDIS seemed confident with the product unlike novice users who only recently started using the system.

#### 4.1.8 Sub-task analysis

Sub-task analysis was also instrumental in identifying a number of usability concerns, including

1. Uneconomical space-usage, dense and scattered data entry sections where it was difficult to quickly identify points of important information (e.g. task 4: Nursing assessment)
2. Lack of auto-population, which forces the nurse to waste time filling out redundant fields throughout the clinical tasks,
3. Ambiguous and nonfunctional data entry requirements, which allow inconsistent documentation throughout clinical tasks.

For example, one expert nurse neglected to enter 'operating room (OR)' as the surgery destination and instead just entered room number where surgery was consulted, but still was able to proceed with the task 9: complete admission pathway without any warning. Similarly, another expert nurse chose "Med-Surgical unit" after struggling with finding OR in the end. These examples revealed that the data selection pathway was not always clear and that it was easy to skip necessary (required) steps in the process and still not be warned of errors, which could lead to delays in care outside of the usability lab.

Similarly, as one nurse documented wound care, she checked the amount of anesthesia but left without selecting the actual value under 'amount' without any error message. The most common and serious usability issue observed for both scenarios regardless of the expertise level was unusable search and navigation functionalities across the clinical tasks. Most nurses were unsuccessful finding the documentation section once they encountered unfamiliar tasks. According to the discussion with the participants in the debriefing sessions, they spent significant time going back and forth and ended up using free text without further attempting to search for appropriate structured data entry templates, because the populated list from which to select a template was too long and insensitive to the query. For example, an expert nurse had trouble finding the right section as she tried to document 'pelvic exam' and ended up typing all information in plain text in a generic nursing note template. Additional usability concerns include non-intuitive names for labels and functions, and the exclusive use of uppercase letters in dropdown menu options, which can decrease the speed of nurses reading through the options.

## 5. Discussion

It has previously been shown that the adoption of HIT, particularly in the interruptive setting of an emergency department, has the potential to enhance the efficiency [19] and quality of medical care by expediting and simplifying clinical workflow while minimizing the effect of human error [42, 43]. This usability evaluation was instrumental in identifying different levels of usability issues. Features and functionalities that should be improved or integrated into the system were recommended in order to increase efficiency and improve the quality of patient care.

### 5.1 Answers to the study questions

In this study, some level of usability gaps existed between expert and novice nurses, as novice nurses completed the tasks both less efficiently and less effectively, and expressed less satisfaction over the EDIS. The most interesting finding in this study was the result of 'percent task success rate,' the most clear performance measure that reflected an accuracy of task, as no substantial difference was observed between two nurse groups. This was entirely contrary to our hypothesis and may suggest that both nurse groups suffer from low usability regardless of experience with the system use. The results may also indicate that both nurse groups spent varying degrees of effort to achieve the same level of successful documentation, which has implications for future clinician training. Why some novices were successful while experts failed in the task 8 (Record follow up/patient's response to medications given) was not certain and further investigation may be needed with more participants to better understand this.



Correlation among performance measures and the task completion survey was contributory in identifying the association among different measures. Overall, the correlation results were mixed or statistically insignificant, while there were a few strong associations found between measures such as time on task vs. mouse movements for the expert group and mouse movements and mouse clicks for the novice group.

## 5.2 Results in relation to other studies

Some of the usability issues elicited in this study have been reported in other EHR usability studies. A usability evaluation study on Armed Force Health Longitudinal Technology Application (AHLTA) EHR by 17 clinicians (14 physicians and 3 nurses) in a laboratory setting revealed several usability concerns, including difficulties in the use of structured documentation as our nurses did [44]. A heuristic usability evaluation by two usability experts on computerized provider order entry (CPOE) interface reported numerous similar usability concerns as ours, including confusing functionalities, dense information display, unusable search function [45]. Similarly, a heuristic walkthrough (HW) and heuristic evaluation of pediatric hospital EHR also elicited numerous usability issues on order functions, navigation, and layout [46].

## 5.3 Weakness of the study

There were some important methodological limitations in this study. This study includes a usability test with a limited number of ED nurses in a single institution as they completed a limited number of clinical tasks. Thus, the clinical task may not encompass or represent other functionalities used in other clinical cases. This study was conducted in a laboratory setting in which a simulated environment cannot account for certain environmental factors in a busy ED, like crowding and interruptions, that may differentially affect the performance of novices vs. experts. Nurses who are familiar with the system may customize his or her interface (e.g. by creating macros and quick lists) to a varying degree, which is not reflected in this study and may have influenced the results. General job inexperience among novice nurses could have affected certain results, such as time-on-task, but because there was no equivalent „paper“ environment also in use by both groups, there was no good way to control for this possible effect. Usability software cannot capture all human cognitive processes as the participant interacts with the application. Currently, the best method to assess awareness may be the use of eye-tracking technology, in which relatively accurate and meaningful behavioral and psychological metrics may be collected [47]. Thus, future studies may be warranted with more nurse evaluators and utilizing more diverse usability evaluation tools.

## 6. Conclusion

We conclude that employing the four sets of quantitative, qualitative usability data collection measures and sub-task analysis was instrumental in identifying the level of usability gaps between expert and novice nurses. In order to obtain significant results, extended studies will incorporate more clinician evaluators. While it may prove difficult to change the current EDIS, we believe this pilot study will serve as a baseline study for a future comparative usability evaluation of EDIS in other institutions with similar clinical settings. These usability study methods may apply to the evaluation of other clinical applications as well.

### Clinical Relevance Statement

This pilot study employed widely accepted usability testing methods in a laboratory setting as ten representative emergency department nurses perform two sets of nine comprehensive scenario-based tasks. The result may also enable developers of clinical information systems to design user-centered EHRs that can promote effective and efficient practitioner performance that leads to the quality of patient care in the end.

**Conflict of Interest**

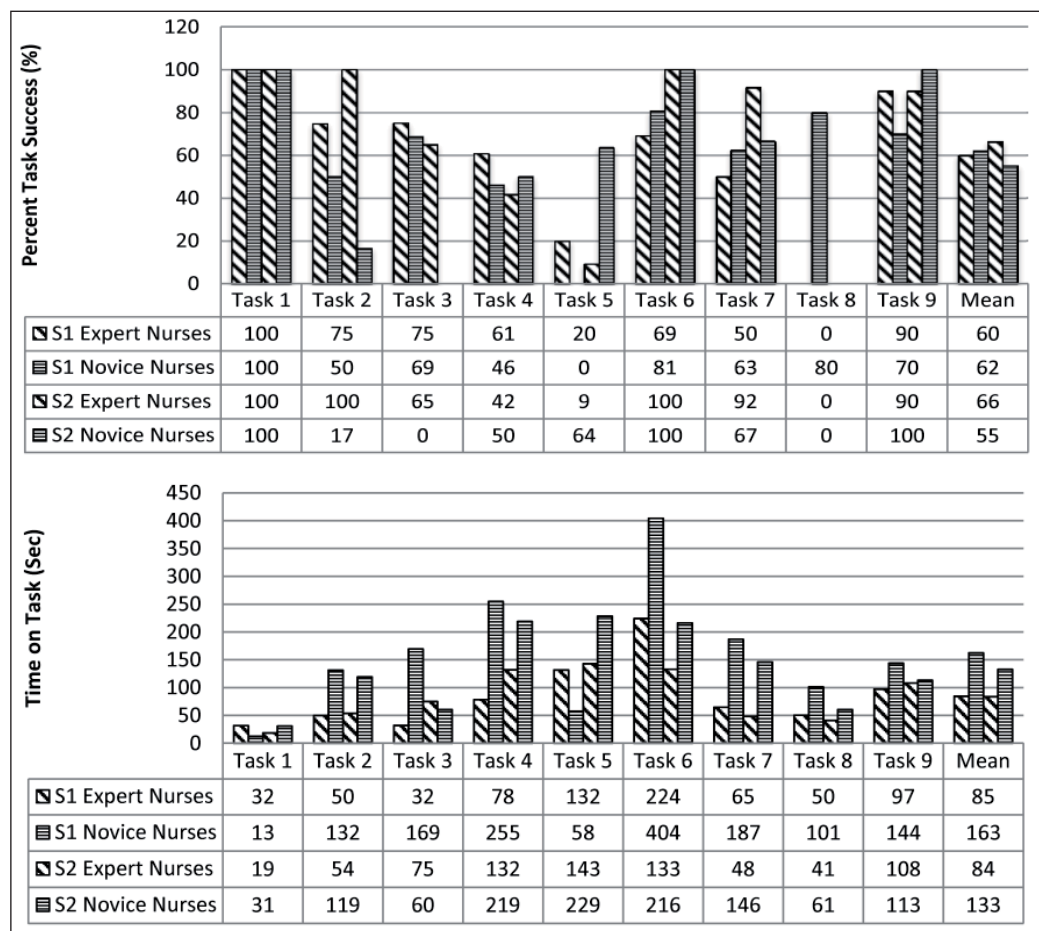
The authors of this paper do not have any financial and personal relationships with other people or organizations that could inappropriately influence this work.

**Human Subjects Protections**

This investigation was approved for human subjects review by the Mount Sinai Institutional Review Board.

**Acknowledgments**

The authors would like to thank Dwayne Raymond, ED Nurse Manager and Robert Asselta, ED Education Specialist from Mount Sinai Medical Center, NYC, NY for participant recruitment, ED nurses for participation of the study, and Michael Badia, IT Clinical Applications Director for technical assistance. At the time of the study, Dr. Shapiro had support through a grant from the National Library of Medicine (5R00LM009556-05).



**Figure 1** Percent task success rate (top) and Time on task (bottom) of nine tasks completed by expert and novice nurse groups. While percent task success rate results shows that there observed no substantial difference in completing tasks, novice nurses spent 1.5–2 times more than expert nurse group in time on task. S1 indicates scenario 1; S2 indicates scenario 2.

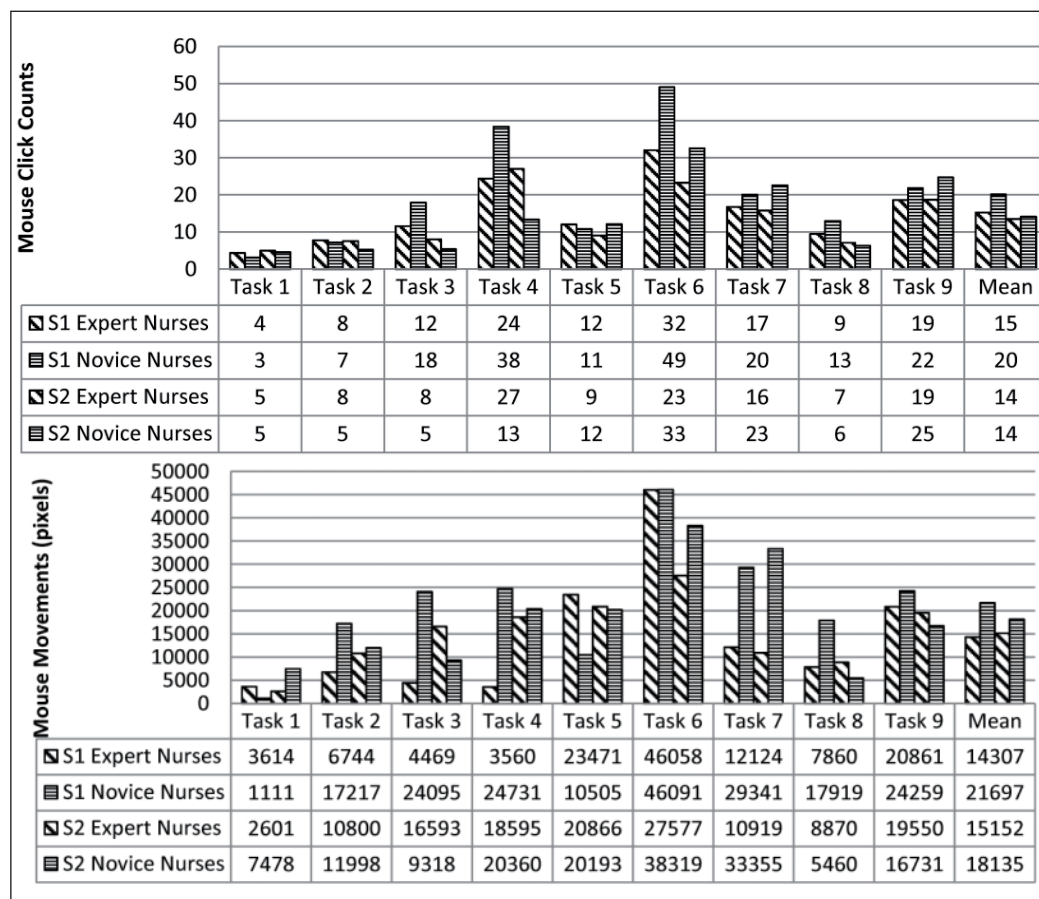


Figure 2 Mouse clicks and mouse movements of nine tasks completed by expert and novice nurse groups. While there observed no substantial difference in mouse clicks between two nurse groups of different expertise levels, expert nurse group showed slightly lower mouse movements across the tasks for both scenarios. S1 indicates scenario 1; S2 indicates scenario 2.

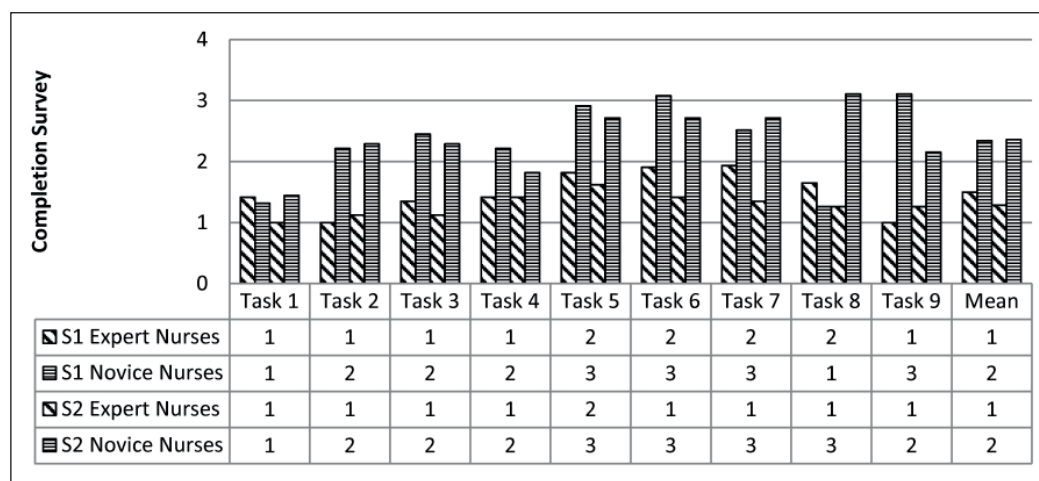


Figure 3 Task completion survey ratings of nine tasks completed by expert and novice nurse groups. There observed no substantial difference in completing tasks between two nurse groups of different expertise levels. S1 indicates scenario 1; S2 indicates scenario 2.

**Table 1** ED scenario of appendicitis case where the patient is ultimately admitted and taken to the operating room (OR). Nine clinical task list of appendicitis case commonly performed by primary nurses. Tasks were coded by task sequences and grouped by functions.

Jane Smith is a 25 year old female who presents to the ED at 8am with a complaint of RLQ abdominal pain. Her triage vitals are HR 80, RR 16, BP 126/70, T (tympanic) 98.9. She is on Allegra for seasonal allergies, has an allergy to penicillin, had a history of tonsillectomy as a child and has no other past medical history. She smokes cigarettes, drinks socially, denies any other drug use and lives with her family. The pain started the night before the periumbilica region, and then localized to her RLQ after several hours. She describes the pain as being sharp with moderate severity. The pain was accompanied by nausea and several episodes of vomiting, and she has persistent anorexia. Abdominal exam reveals tenderness at McBurney's point and a positive Rovsing's sign, the abdomen is otherwise flat, non-distended and non-tender with no peritoneal signs. A pelvic exam is done and is unremarkable, as is the rest of the physical. Bedside urinalysis (UA) and urine pregnancy tests are ordered. Labs are ordered including a CBC, chemistry panel, liver function test (LFT). A rectal temperature and a contrast CT of the abdomen and pelvis are also ordered. Consent is obtained for the procedure. Bedside urinalysis (UA) and urine pregnancy tests are negative. Rectal temp is 100.4, the patient has as WBC of 14.2, and the patient CT test is positive for appendicitis; labs are otherwise normal. Antibiotics are ordered (Levaquin and Flagyl) and administered intravenously (IV). Surgery is consulted and the patient is ultimately admitted to their service and then taken to the operating room (OR).

Task Code	Task Name	Information Given	Desired Sub-tasks
Task1	Log in	Username: n01, Password: n01	<ol style="list-style-type: none"> <li>1. Click on the Main icon</li> <li>2. Type UN – n01, PW- n01</li> <li>3. Enter</li> </ol>
Task2	Enter allergy/ current meds	Allegra for seasonal allergies, has an allergy to penicillin,	<ol style="list-style-type: none"> <li>1. Go to Current Medication</li> <li>2. Search – Select Allegra</li> <li>3. Click Enter/Allergy</li> <li>4. Check Environmental (Seasonal)</li> <li>5. Search – Select Penicillin</li> <li>6. Enter</li> </ol>
Task3	Complete PMH	She had a history of tonsillectomy as a child and has no other past medical history. She smokes cigarettes, drinks socially, denies any other drug use and lives with her family.	<ol style="list-style-type: none"> <li>1. Go to Past Medical History</li> <li>2. Medical History – Check No PMH</li> <li>3. Surgical History – Check tonsillectomy</li> <li>4. Social History – Check Denies Drug abuse</li> <li>5. Social History – Choose Smokes / Tobacco</li> <li>6. Social History – Choose Alcohol / Socially</li> <li>7. Social History – Check Housing status – Home family</li> <li>8. Enter</li> </ol>
Task4	Nursing assessment	The pain started the night before periumbilically, and then localized to her RLQ after several hours. She describes the pain as being sharp with moderate severity. The pain was accompanied by nausea and several episodes of vomiting, and she has persistent anorexia. Abdominal exam reveals tenderness at McBurney's point and a + Rovsing's sign, the abdomen is otherwise flat, non-distended and non-tender with no peritoneal signs.	<ol style="list-style-type: none"> <li>1. Go to Assess</li> <li>2. Go to Nursing Abdomen</li> <li>3. Abdomen – Choose Pain Site RLQ</li> <li>4. Abdomen – Choose Pain Quality Sharp</li> <li>5. Abdomen – Choose Pain Scale Moderate (5)</li> <li>6. Abdomen – Check Nausea Present</li> <li>7. Abdomen – Choose Vomiting</li> <li>8. Abdomen – Document Anorexia</li> <li>9. Abdomen – Document tenderness at McBurney's point and a + Rovsing's sign</li> <li>10. Abdomen – Check nontender</li> <li>11. Abdomen – Check nondistended</li> <li>12. Abdomen – Check no peritoneal signs</li> <li>13. Enter</li> </ol>

**Table 1** Continued

Task Code	Task Name	Information Given	Desired Sub-tasks
Task5	Document procedure with option to enter vital signs	A pelvic exam is done and is unremarkable, as is the rest of the physical.	<ol style="list-style-type: none"> <li>1. Go to RN Procedures</li> <li>2. Pelvic Exam – Document Time</li> <li>3. Pelvic Exam – Document Procedure</li> <li>4. Pelvic Exam – Document Safety</li> <li>5. Enter</li> </ol>
Task6	Order diagnostic tests	Bedside UA and urine pregnancy tests are ordered. Labs are ordered including a CBC, chem 7, LFT's. A rectal temperature, and a Contrast CT of the abdomen and pelvis are also ordered and the patient is consented for the procedure.	<ol style="list-style-type: none"> <li>1. Go to Orders</li> <li>2. Check Urinalysis</li> <li>3. Urine Pregnancy Test</li> <li>4. Check CBC, PLT, and DIFF</li> <li>5. Check ER Venous Panel</li> <li>6. Check Abdominal Pain Panel</li> <li>7. Check ICON</li> <li>8. Check Urinalysis</li> <li>9. Custom order – Rectal temperature</li> <li>10. Choose Ordering Physician</li> <li>11. Enter</li> <li>12. Common Radiology – CT Scan -Abdomen/Pelvis w/ contrast</li> <li>13. Enter</li> </ol>
Task7	Document results of tests	Bedside UA and urine pregnancy tests are negative. Rectal temp is 100.4, the patient has as WBC of 14.2, and the patient CT tests positive for appendicitis; labs are otherwise normal.	<ol style="list-style-type: none"> <li>1. Go to RN Procedure</li> <li>2. Bedside testing – Pregnancy test – negative</li> <li>3. Bedside testing – Urine Dip -negative</li> <li>4. Rectal temp – 100.4</li> <li>5. WBC – 14.2</li> <li>6. CT – positive for appendicitis</li> <li>7. Mark other labs negative</li> <li>8. Enter</li> </ol>
Task8	Record follow up/patient's response to medications given	No reactions were observed	<ol style="list-style-type: none"> <li>1. Go to Medication Service</li> <li>2. Click Follow up – Levaquin – document no adverse effect</li> <li>3. Enter</li> <li>4. Click Follow up – Metronidazole- document no adverse effect</li> <li>5. Enter</li> </ol>
Task9	Complete admission pathway	Surgery is consulted and the patient is ultimately admitted to their service and then taken to the OR.	<ol style="list-style-type: none"> <li>1. Go to RN Procedure</li> <li>2. Admission – Time</li> <li>3. Admission – Procedure – Patient admitted to OR</li> <li>4. Admission – Transportation</li> <li>5. Enter</li> </ol>

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

**Table 2** ED scenario of ambulatory condition requiring a minor procedure. The patient is discharged with instruction after procedures. Nine clinical task lists of minor procedure case, which were commonly performed by primary nurses. The tasks were coded by task sequences and grouped by functions.

John Smith is a right-handed 39 year old patient with a history of diabetes who accidentally put his hand through a glass window about 2 hours prior to arrival sustaining a 2 cm laceration to the dorsum of his left hand. His triage vitals are HR 76, RR 16, BP 132/80, T 98.6. He is on metformin for his diabetes, has no other medical problems or medications, no known drug allergies (NKDA) and has no past surgical history (PSH). He lives alone and denies drug, alcohol or tobacco use. He does not remember his last tetanus shot. He is placed in urgent care. On exam, the patient has no loss of strength and distal sensation is intact. There is no visible tendon injury or foreign body. Finger stick (FS) is ordered for glucose level. An x- ray is ordered to rule out glass foreign body. His FS glucose is 207 and x-ray is negative. A tetanus shot is also ordered and administered by the nurse. The wound is infiltrated with ~ 5 cc of 2% lidocaine with epinephrine and copiously irrigated using a Zerowet and 60 cc of sterile water. The wound is then sutured using 4–0 nylon sutures and 4 interrupted stitches, then dressed. The patient tolerated the procedure well. The patient is given discharge (d/c) instructions and told to return in 48 hours for a wound check and in 7 days for suture removal. Because of his diabetes, he is given a prescription for Keflex.

Task Code	Task Name	Information Given	Desired Sub-tasks
Task1	Log in	Username: n01, Password: n01	<ol style="list-style-type: none"> <li>1. Click on the Main icon</li> <li>2. Type UN – n01, PW – n01</li> <li>3. Enter</li> </ol>
Task2	Enter allergy/ current meds	He is on metformin for his diabetes, has no other medical problems or medications, NKDA (no known drug allergies)	<ol style="list-style-type: none"> <li>1. Go to Current Medication</li> <li>2. Search Metformin</li> <li>3. Select Metformin Hydrochloride ER</li> <li>4. Click Enter/Allergy</li> <li>5. Check No Known Drug Allergies</li> <li>6. Enter</li> </ol>
Task3	Complete PMH	Has no PSH. He lives alone and denies drug, alcohol or tobacco use. He does not remember his last tetanus shot.	<ol style="list-style-type: none"> <li>1. Go to Past Medical History</li> <li>2. Medical History – Check Diabetes</li> <li>3. Medical History – Check Diabetes – on oral medication</li> <li>4. Surgical History – No surgical history</li> <li>5. Social History – Check Denies alcohol abuse</li> <li>6. Social History – Check Denies Tobacco abuse</li> <li>7. Social History – Check Denies drug abuse</li> <li>8. Social History – Check Lives alone</li> </ol>
Task4	Nursing assessment	He accidentally put his hand through a glass window about 2 hours prior to arrival sustaining a 2 cm laceration to the dorsum of his left hand. On exam, the patient has no loss of strength and distal sensation is intact. There is no visible tendon injury or foreign body.	<ol style="list-style-type: none"> <li>1. Go to Assess</li> <li>2. Go to Nursing Extremity Upper</li> <li>3. Left Upper Extremity – Check Sensation intact</li> <li>4. Left Upper Extremity – Note – document L hand laceration</li> <li>5. Note – no visible tendon injury or foreign body</li> <li>6. Enter</li> </ol>

**Table 2** Continued

Task Code	Task Name	Information Given	Desired Sub-tasks
Task5	Document procedure with option to enter vital signs	The wound is infiltrated with ~ 5 cc of 2% lidocaine with epinephrine and copiously irrigated using a zerowet and 60 cc of sterile water. The wound is then sutured using 4.0 ethilon sutures and interrupted stitches x 4, then dressed.	<ol style="list-style-type: none"> <li>1. Go to RN Procedures</li> <li>2. Wound Care – Document Time</li> <li>3. Wound Care – Site – L hand</li> <li>4. Wound Care – Anesthesia – Anesthetic used</li> <li>5. Wound Care – Anesthesia – Check Medication used – 2% Lidocaine with epinephrine</li> <li>6. Wound Care – Anesthesia – Amount – 5 cc</li> <li>7. Wound Care – Wound Irrigation – Wound irrigated – Other amount – with Zerowet and 60 cc of sterile water</li> <li>8. Wound Care – Wound repair – Wound repaired with – sutures</li> <li>9. Wound Care – Notes – The wound is then sutured using 4.0 ethilon sutures and 4 interrupted stitches</li> <li>10. Wound Care – Dressing -</li> <li>11. Enter</li> </ol>
Task6	Order diagnostic tests	Finger stick is ordered for glucose level. An x ray is ordered to rule out glass foreign body.	<ol style="list-style-type: none"> <li>1. Go to Orders</li> <li>2. Common Radiology – X-ray – Left hand</li> <li>3. Common Radiology – X-ray – Reason for test – r/o foreign body</li> <li>4. Common Radiology – X-ray – Relevant History – laceration via glass</li> <li>5. Common Radiology – X-ray -Is patient pregnant? – No</li> <li>6. Choose Ordering Physician</li> <li>7. Enter</li> <li>8. Nursing – Finger stick</li> <li>9. Choose Ordering Physician</li> <li>10. Enter</li> </ol>
Task7	Document results of tests	His FS glucose is 207 and x ray is negative.	<ol style="list-style-type: none"> <li>1. Go to RN Procedure</li> <li>2. Bedside testing – Glucose – Indications – Diabetic patient</li> <li>3. Bedside testing – Glucose – Procedure – blood – capillary</li> <li>4. Bedside testing – Glucose – Procedure – results – 207</li> <li>5. Nurse Notes – X-ray test negative</li> <li>6. Enter</li> </ol>
Task8	Record follow up/patient's response to medications given	No reactions were observed	<ol style="list-style-type: none"> <li>1. Go to Medication Service</li> <li>2. Click Follow up – Tetanus – document no adverse effect</li> <li>3. Enter</li> </ol>
Task9	Complete discharge pathway	The patient is given d/c instructions and told to return in 48 hours for a wound check and in 7 days for suture removal. Because of his diabetes, he is given a prescription for Keflex	<ol style="list-style-type: none"> <li>1. Go to RN Procedure</li> <li>2. Discharge Note – Special instructions – Discharge instructions given to -Patient</li> <li>3. Discharge Note – Special instructions – Name of prescription given – Keflex</li> <li>4. Discharge Note – Notes – told to return in 48 hours for a wound check and in 7 days for suture removal</li> <li>5. Enter</li> </ol>



**Table 3** Pearson correlation coefficients among performance measures and subjective task completion survey ratings by expert (N = 108) and novice nurse groups (N = 63). As the interpretation guideline, correlation coefficients between [0, 0.3) are considered weak, those between [0.3 and 0.7) are moderate and coefficients between [0.7 and 1] are considered high. “[” indicates inclusive and “)” indicates exclusive. Acronyms were used for succinct representation of the measures in the table. MC: Mouse Clicks, MM: Mouse Movements, TOT: Time on Task, TCS: Task Completion Survey, PSR: Percent Task Success Rate. Values with bold type indicate strong correlation.

		PTSR	MC	MM	TOT
MC	Expert	-0.05 ( $p = 0.60$ )			
	Novice	0.19 ( $p = 0.14$ )			
MM	Expert	-0.01 ( $p = 0.92$ )	0.20 ( $p = 0.04$ )		
	Novice	0.14 ( $p = 0.27$ )	<b>0.64</b> ( $p = 0.00$ )		
TOT	Expert	0.01 ( $p = 0.92$ )	0.30 ( $p = 0.00$ )	<b>0.83</b> ( $p = 0.00$ )	
	Novice	0.04 ( $p = 0.76$ )	<b>0.68</b> ( $p = 0.00$ )	<b>0.74</b> ( $p = 0.00$ )	
TCS	Expert	-0.30 ( $p = 0.76$ )	0.04 ( $p = 0.68$ )	0.21 ( $p = 0.03$ )	0.40 ( $p = 0.00$ )
	Novice	-0.35 ( $p = 0.00$ )	0.05 ( $p = 0.70$ )	0.24 ( $p = 0.06$ )	0.32 ( $p = 0.01$ )

## References

1. Gans D, Kralewski J, Hammons T, Dowd B. Medical groups' adoption of electronic health records and information systems. *Health Aff (Millwood)* 2005; 24(5): 1323–1333.
2. Bertman J, Skolnik N. Poor usability keeps EHR adoption rates low. *Family Practice News* 2010; May 1.
3. Technology NIOsa. Common industry specification for usability requirements NISTIR 7432. Gaithersburg, MD 2007. Report No.: IUSR:CISU-R v0.90.
4. Kim MS, Mohrer D, Trusko B, Landrigan P, Elkin P. World Trade Center Medical Monitoring and Treatment Program: A clinical workflow analysis. *AMIA Clinical Research Informatics Summit 2010*; San Francisco, CA 2010.
5. Beuscart-Zephir MC, Elkin P, Pelayo S, Beuscart R. The human factors engineering approach to biomedical informatics projects: state of the art, results, benefits and challenges. *Yearb Med Inform* 2007; 109–127.
6. Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, Strom BL. Role of computerized physician order entry systems in facilitating medication errors. *JAMA* 2005; 293(10): 1197–1203.
7. Armijo D, McDonnell C, Werner K. Electronic health record usability: Interface design considerations. *AHRQ Publication No. 09(10)-0091–2-EF*. Rockville, MD: Agency for Healthcare Research and Quality 2009.
8. Armijo D, McDonnell C, Werner K. Electronic health record usability-evaluation and use case framework. *AHRQ Publication No. 09(10)-0091–1-EF*. Rockville, MD: Agency for Healthcare Research and Quality 2009.
9. Belden JL, Grayson R, Barnes J. Defining and testing EMR usability: Principles and proposed methods of EMR usability evaluation and rating. *HIMSS EHR Usability Task Force* 2009.
10. Kane LR. Electronic medical record survey results: Medscape exclusive readers' choice. *Medscape Business of Medicine* 2009.
11. Edsall RL, Adler KG. The 2009 EHR user satisfaction survey responses from 2,012 family physicians. *Family Practice Management* 2009; 16(6).
12. McDonnell C, Werner K, Wendel L. Electronic health record usability vendor practices and perspectives. Rockville, MD: Agency for Healthcare Research and Quality 2010.
13. Smelcer J, Miller-Jacobs H, Kantrovich L. Usability of electronic medical records. *Journal of Usability Studies* 2009; 4(2): 70–84.
14. Current State of Emergency Department Information Systems. *InfoHealth Management Corp* 2007.
15. Pallin D, Lahman M, Baumlin K. Information technology in emergency medicine residency-affiliated emergency departments. *Acad Emerg Med* 2003; 10(8): 848–852.
16. Taylor TB. Information management in the emergency department. *Emerg Med Clin North Am* 2004; 22(1): 241–257.
17. Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, Morton SC, Shekelle PG. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med* 2006 May; 144(10): 742–752.
18. Kaushal R, Barker KN, Bates DW. How can information technology improve patient safety and reduce medication errors in children's health care? *Arch Pediatr Adolesc Med* 2001; 155(9): 1002–1007.
19. Baumlin KM, Shapiro JS, Weiner C, Gottlieb B, Chawla N, Richardson LD. Clinical information system and process redesign improves emergency department efficiency. *Jt Comm J Qual Patient Saf* 2010; 36(4): 179–185.
20. Shapiro JS, Baumlin KM, Chawla N, Genes N, Godbold J, Ye F, Richardson LD. Emergency department information system implementation and process redesign result in rapid and sustained financial enhancement at a large academic center. *Acad Emerg Med* 2010; 17(5): 527–535.
21. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? *Hum Factors* 1992; 34(4): 457–468.
22. Tullis T, Albert B. Issues-based metrics. Measuring the user experience. San Francisco: Morgan Kaufmann; 2008; 99–121.
23. Barnum C. The magic number 5: Is it enough for web testing? *Information Design Journal* 2003; 11: 160–170.
24. Lewis JR. Evaluation of procedures for adjusting problem-discovery rates estimated from small samples. *The International Journal of Human-Computer Interaction* 1994; 13(4): 445–479.
25. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*; Amsterdam, The Netherlands. 169166: ACM; 1993; 206–213.
26. Lewis JR. Sample sizes for usability studies: additional considerations. *Human Factors* 1994; 36(2): 368–378.

27. Sauro J, Dumas JS. Comparison of three one-question, post-task usability questionnaires. CHI 2009; Boston, MA2009.
28. Brooke J. SUS—a quick and dirty usability scale. Usability Evaluation in Industry. London ; Bristol, Pa.: Taylor & Francis; 1996.
29. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction* 2008; 24(6): 574–594.
30. Lewis JR, Sauro J. The factor structure of the system usability scale. *Proceedings of the 1st International Conference on Human Centered Design: Held as Part of HCI International 2009; San Diego, CA.* 1601645: Springer-Verlag; 2009; 94–103.
31. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies* 2009; 4(3): 114–123.
32. Lutes KD, Chang K, Baggili IM, editors. Diabetic e-Management System (DEMS). *Information Technology: New Generations, 2006 ITNG 2006 Third International Conference on;* 2006 10–12 April 2006.
33. Kastner M, Lottridge D, Marquez C, Newton D, Straus SE. Usability evaluation of a clinical decision support tool for osteoporosis disease management. *Implementation science* : IS2010; 5: 96.
34. Suominen O, Hyvonen E, Viljanen K, Hukka E. Health Finland – a national semantic publishing network and portal for health information. *Web Semant* 2009; 7(4): 287–297.
35. CCHIT 2011 Usability Testing Guide for Ambulatory EHR's. *Certification Commission for Health Information Technolog* 2009.
36. Schumacher RM, Lowry SZ. NIST guide to the processes approach for improving the usability of electronic health records. *National Institute of Standards and Technology, U.S. Department of Commerce* 2010 Contract No.: NISTIR 7741.
37. Cordes RE. The effects of running fewer subjects on time-on-task measures. *International Journal of Human-Computer Interaction*1993; 5(4): 393–403.
38. Eisenhart C, Deming L, Martin CS. On the arithmetic mean and median in small samples from the normal and certain non-normal populations. *Annals of Mathematical Statistics* 1948: 599.
39. Anschuetz L. How does corporate culture affect the reporting of usability results? UPA 2004 Idea Markets [serial on the Internet]. 2004.
40. Nielsen J. First rule of usability? Don't listen to users. 2001.
41. Sauro J. Do users fail a task and still rate it as easy? 2009; Available from: <http://www.measuringusability.com/failed-sat.php>.
42. Kuperman GJ, Teich JM, Gandhi TK, Bates DW. Patient safety and computerized medication ordering at Brigham and Women's Hospital. *Joint Commission Journal on Quality and Patient Safety* 2001; 27(10): 509–521.
43. Buller-Close K, Schriger DL, Baraff LJ. Heterogeneous effect of an emergency department expert charting system. *Ann Emerg Med* 2003; 41(5): 644–652.
44. Stagers N, Jennings BM, Lasome CEM. A usability assessment of AHLTA in ambulatory clinics at a military medical center. *Military Medicine* 2010; 175(7): 518–524.
45. Li Q, Douglasa S, Hundt AS, Carayona P, editors. A heuristic usability evaluation of a computerized provider order entry (CPOE) technology. *IEA2006 Congress*; 2006.
46. Edwards PJ, Moloney KP, Jacko JA, Sainfort F. Evaluating usability of a commercial electronic health record: A case study. *International Journal of Human-Computer Studies* 2008; 66(10): 718–728.
47. Nielsen J, Pernice K. *Eyetracking web usability.* Berkeley, CA.: New Riders; 2010.