

Safeuristics! Do Heuristic Evaluation Violation Severity Ratings Correlate with Patient Safety Severity Ratings for a Native Electronic Health Record Mobile Application?

Brandan Kennedy^{1,2,3} Ellen Kerns⁴ Y. Raymond Chan^{2,3} Barbara S. Chaparro⁵ Sarah D. Fouquet³

¹Medical Informatics and Telemedicine, Children's Mercy Hospital, Kansas City, Missouri, United States

²Pediatric Hospital Medicine, Children's Mercy Hospital, Kansas City, Missouri, United States

³Human Factors Collaborative, Children's Mercy Hospital, Kansas City, Missouri, United States

⁴Department of Pediatrics, University of Nebraska Medical Center, Omaha, Nebraska, United States

⁵Human Factors and Behavioral Neurobiology, Embry-Riddle Aeronautical University, Daytona Beach, Florida, United States

Address for correspondence Brandan Kennedy, MD, Medical Informatics and Telemedicine/Pediatric Hospital Medicine, Children's Mercy Hospital, 2401 Gillham Road, Kansas City, MO 64108, United States (e-mail: b1kennedy@cmh.edu).

Appl Clin Inform 2019;10:210–218.

Abstract

Objective Usability of electronic health records (EHRs) remains challenging, and poor EHR design has patient safety implications. Heuristic evaluation detects usability issues that can be classified by severity. The National Institute of Standards and Technology provides a safety scale for EHR usability. Our objectives were to investigate the relationship between heuristic severity ratings and safety scale ratings in an effort to analyze EHR safety.

Materials and Methods Heuristic evaluation was conducted on seven common mobile EHR tasks, revealing 58 heuristic violations and 28 unique usability issues. Each usability issue was independently scored for severity by trained hospitalists and a Human Factors researcher and for safety severity by two physician informaticists and two clinical safety professionals.

Results Results demonstrated a positive correlation between heuristic severity and safety severity ratings. Regression analysis demonstrated that 49% of safety risk variability by clinical safety professionals ($r = 0.70$; $n = 28$) and 42% of safety risk variability by clinical informatics specialists ($r = 0.65$; $n = 28$) was explained by usability severity scoring of problems outlined by heuristic evaluation. Higher severity ratings of the usability issues were associated with increased perceptions of patient safety risk.

Discussion This study demonstrated the use of heuristic evaluation as a technique to quickly identify usability problems in an EHR that could lead to safety issues. Detection of higher severity ratings could help prioritize failures in EHR design that more urgently require design changes. This approach is a cost-effective technique for improving usability while impacting patient safety.

Conclusion Results from this study demonstrate the efficacy of the heuristic evaluation technique to identify usability problems that impact safety of the EHR. Also, the use of interdisciplinary teams for evaluation should be considered for severity assessment.

Keywords

- ▶ user–computer interface
- ▶ electronic health records
- ▶ mobile application
- ▶ heuristics
- ▶ safety

received
July 26, 2018
accepted after revision
January 31, 2019

© 2019 Georg Thieme Verlag KG
Stuttgart · New York

DOI <https://doi.org/10.1055/s-0039-1681073>.
ISSN 1869-0327.

Background and Significance

Over the past 15 years, health care in the United States has evolved with the rapid deployment of electronic health records (EHRs).¹ The pace of this change has primarily been driven by federal regulations and stimulus funds² with the aim of improving the quality and safety of health care.³ Unfortunately, the pressure for the rapid development and deployment of EHR by EHR vendors often leads to suboptimal designs, ultimately resulting in poor interface usability and difficulties with integration into existing workflow.⁴ In an attempt to address these concerns, the Office of the National Coordinator for Health Information Technology (ONC) published regulations in 2014 entitled “Safety-enhanced Design” within its EHR certification requirements. These regulations focused on user-centered design and usability testing in an effort to drive improvement in EHR workflows, efficiency, and patient safety.⁵ User-centered design is a process framework in which representative end users are engaged early and often throughout the design and development process.⁶ Interviews and observation techniques are used to understand end-user requirements and conceptual models before design work begins.⁷ Usability testing is a technique in which representative users are observed as they complete tasks with a software application or prototype in a controlled setting. Objective and subjective data are collected to determine areas where the application works well or needs improvement. Usability testing has been shown to be an effective technique for discovering usability issues in medical applications.⁸

Another popular method that is used to detect usability issues is heuristic evaluation.⁹ Heuristic evaluation was first described by Nielsen and Molich¹⁰ as an “informal method of usability analysis” and is defined as being more cost-effective than usability testing.¹¹ Heuristic evaluations may be used to capture usability issues, though the number of issues found varies depending on the number of evaluators.¹¹ In addition, heuristic evaluation is not viewed as a substitute for usability testing.¹² To execute this technique, a software interface is assessed by trained evaluators based on 10 design rules, or heuristics, for a set of representative user tasks. For every interaction, user decision, and prompt, reviewers designate whether each heuristic is either obeyed or violated.¹³ The heuristic violations are grouped into usability issues and then rated for severity, which helps determine whether the violation is minor (e.g., aesthetic) or catastrophic (e.g., serious redesign required). Although heuristics can be used to establish severity of design flaws, they do not specifically address how violations might impact patient safety when used in health care. Zhang et al applied Nielsen’s heuristic evaluation to medical devices and expanded the number of heuristics from 10 to 14.^{14,15} Zhang et al reasoned that since a majority of safety issues arose from user error and not device failure,¹⁶ it was important to identify usability issues that could, in turn, reduce the potential for safety events. Subsequently, Zhang et al’s heuristics have been used to evaluate several different medical products including infusion pumps in the intensive care unit,¹⁷ medical information web pages,¹⁸ and telemedicine systems.¹⁹ In addition, researchers have employed portions of Zhang et al’s

heuristics to expedite the heuristic evaluation process even further for a medical information prototype.¹⁸

While some literature suggests that EHRs have improved both quality and safety of health care,^{20,21} significant safety risks in patient care continue to be associated with the use of EHRs.^{22,23} These risks include morbidity and mortality,²⁴ diagnostic errors,²⁵ disruptions in clinical workflow, and negative effects on patient–provider interactions.²⁶ These safety issues have led to a call to be more proactive in preventing errors through better EHR design.^{23,27,28} In 2012, The National Institute of Standards and Technology (NIST) published a document that outlines a safety rating system for EHRs,²⁹ whereas Borycki et al developed an evidence-based safety heuristics methodology for health information systems.³⁰ This study seeks to understand if a relationship exists between the level of severity of heuristic evaluation violations (Nielsen³¹ and Zhang et al¹⁵) and safety risk severity as defined by the NIST safety scale.²⁹

Materials and Methods

To investigate a possible relationship between usability severity scale scores and safety severity scale scores, seven specific provider tasks using a native mobile EHR application that has been commercially available and in use for more than 5 years were evaluated. None of the authors or evaluators in this study was involved in the design or development of the EHR application. The following commonly used tasks within a tertiary pediatric health center were used for the evaluation:

- Order a chest X-ray.
- Cancel the previously ordered chest X-ray.
- View immunization history.
- Start a progress note.
- Locate a specialist note.
- Dictate a progress note.
- Capture a picture of a simulated wound.

The tasks were selected as commonly used tasks for patient care within the mobile application that incorporated different workflows including record review, provider order entry, provider documentation, and media capture. It should be noted that medication ordering, one of the most common tasks in provider workflow, was not available in the mobile application at the time of the study and therefore could not be included in the analyzed tasks.

Procedure

The task evaluation process was divided into two phases. Phase 1 consisted of the heuristic evaluation that identified usability issues and applied severity ratings to each usability issue. This evaluation was then followed by Phase 2, wherein a different group of evaluators rated safety severity (specific to patient safety) for each usability issue identified in Phase 1.

Phase 1: Heuristic Evaluation

Using the Nielsen–Shneiderman heuristics that have been adapted for medical devices (→ **Table 1**),¹⁵ two practicing pediatric hospitalists and one human factors researcher each

Table 1 Nielsen–Shneiderman heuristics for medical devices

Name [abbreviation]	Definition
Consistency and standards [consistency]	Users should not have to wonder whether different words, situations, or actions mean the same thing. Standards and conventions in product design should be followed.
Visibility of system state [visibility]	Users should be informed about what is going on with the system through appropriate feedback and display of information.
Match between system and world [match]	The image of the system perceived by users should match the model the users have about the system.
Minimalist [minimalist]	Any extraneous information is a distraction and a slowdown.
Minimize memory load [memory]	Users should not be required to memorize a lot of information to carry out tasks. Memory load reduces users' capacity to carry out the main tasks.
Informative feedback [feedback]	Users should be given prompt and informative feedback about their actions.
Flexibility and efficiency [flexibility]	Users always learn and are always different. Give users the flexibility of creating customization and shortcuts to accelerate their performance.
Good error messages [message]	The messages should be informative enough such that users can understand the nature of errors, learn from errors, and recover from errors.
Prevent errors [error]	It is always better to design interfaces that prevent errors from happening in the first place.
Clear closure [closure]	Every task has a beginning and an end. Users should be clearly notified about the completion of a task.
Reversible actions [undo]	Users should be allowed to recover from errors. Reversible actions also encourage exploratory learning.
Use users' language [language]	The language should be always presented in a form understandable by the intended users.
Users in control [control]	Do not give users the impression that they are controlled by the system.
Help and documentation [document]	Always provide help when needed.

Source: Adapted from Zhang et al.¹⁵

performed an initial heuristic evaluation of the seven designated tasks.

Prior to the heuristic evaluation, the human factors researcher trained both hospitalists on the method of heuristic evaluation. The training process included practice using several examples from the study by Zhang et al.¹⁴ The hospitalists had more than 6 months of experience using the native EHR mobile application to be evaluated. To better understand the phase 1 method, see [►Fig. 1](#).

The human factors researcher initially performed her own heuristic evaluation of each task in a separate session. Then the hospitalists performed each task in independent sessions with the human factors researcher, who recorded the heuristic evaluation results of the hospitalists, asked clarifying questions, and recorded individual anecdotal comments in each session. The hospitalists were blind to each other's results as well as to the results of the human factors researcher. After the heuristic evaluations were completed and the specific violations were identified, both hospitalists and the human factors researcher held a subsequent consensus review session to define the usability issues corresponding to the heuristic violations from each task. The defined usability issues were then assigned a single severity scale score based on the following three principles from Nielsen:³¹ The *frequency* (rare vs. common) with which the problem occurs, the *impact* (easy or difficult to overcome) of

the problem if it occurs, and the *persistence* (occurs sometimes or always in a task workflow) of the problem. Issues were given a rating from 0 (no problem) to 4 (usability catastrophe) ([►Table 2](#), column 1).

Phase 2: Safety Evaluation

Considering its potential effect on patient safety, usability is a particularly important aspect of EHR design. Usability issues can lead to user error, which, in turn, can have potential safety consequences for patients. Because the heuristic evaluation severity scores are specific to usability and do not take into account issues that may contribute to patient safety risk (e.g., the violation could contribute to patient mortality or morbidity through workarounds or inappropriate actions), the decision was made to recruit independent raters to examine the usability issues strictly on the premise of patient safety using a proposed safety severity scale. The NIST safety scale was developed as part of an EHR usability protocol (NISTIR 7804–1).²⁹ The safety scale evaluation was individually and independently performed by one board-certified pediatrician and clinical informatics specialist, one board-certified pediatric intensivist and clinical informatics specialist, and two clinical safety officers certified by the Board for Professionals in Patient Safety. All received identical training at the beginning of their individual sessions on the NIST safety scale. All were blind to the initial usability

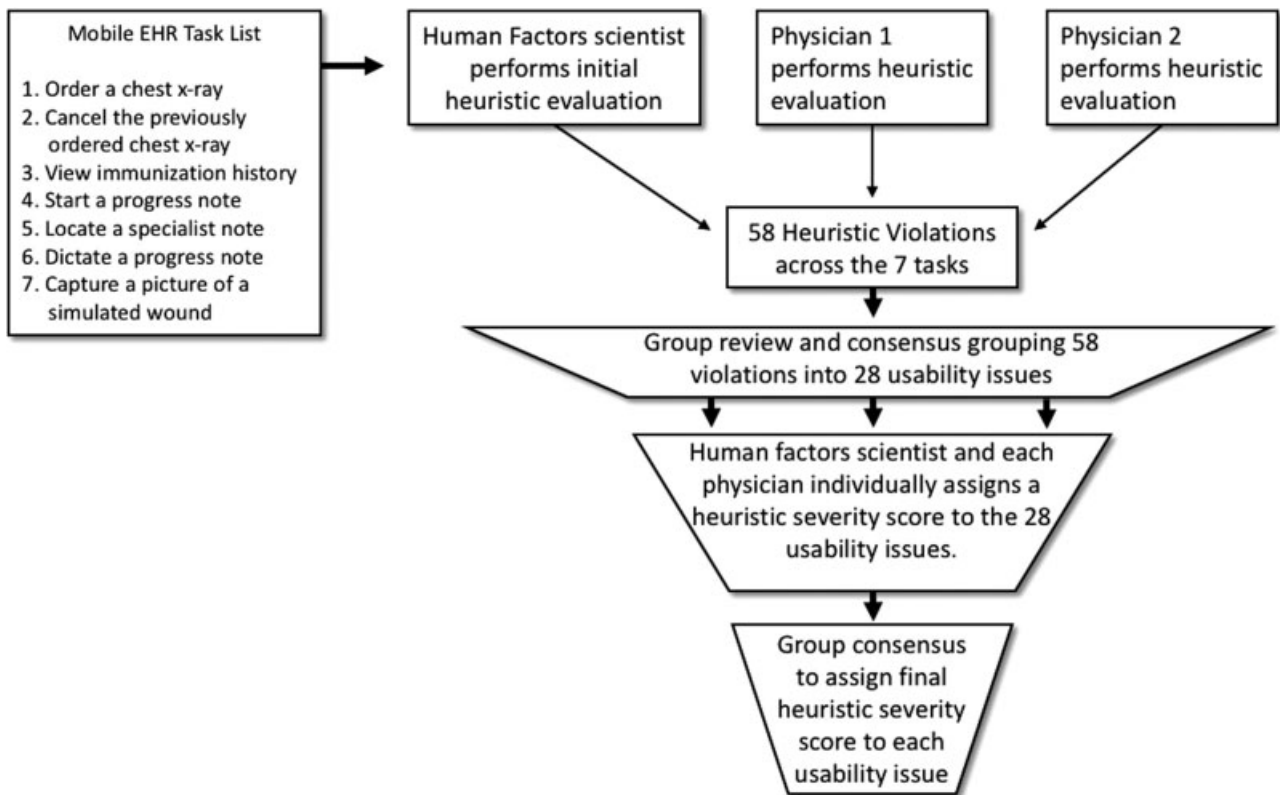


Fig. 1 Phase 1 methods flow diagram.

Table 2 Rating severity scale level definitions by type

Rating	Heuristic (Nielsen ³¹)	Safety (Lowry et al ²⁹)
0	I do not agree that this is a usability problem at all	No issue/not applicable
1	Cosmetic problem only: need not be fixed unless extra time is available on project	Minor: potential for lower quality of clinical care due to decreased efficiency, increased frustration, or increased documentation burden or workload
2	Minor usability problem: fixing this should be given low priority	Moderate: potential for workarounds that create patient safety risks
3	Major usability problem: important to fix and therefore should be given high priority	Major: potential for patient morbidity
4	Usability catastrophe: imperative to fix this before the product can be released	Catastrophic: potential for patient mortality

severity ratings of the heuristic evaluation (neither physician had been involved in the initial heuristic evaluation sessions). Each participant then reviewed each previously identified usability problem and rated the severity according to the safety severity scale (→Table 2, column 2).

The data were analyzed using R statistical software. The combined maximum safety severity score among physician informaticists and the combined maximum safety severity score among clinical safety professionals were used in all summary comparisons. The maximum safety scores were used for this analysis because it aligns with the “speaking up” movement within patient safety. “Speaking up” refers to

encouraging individuals to voice safety concerns regardless of the perceived certainty or popularity related to the concern. Summary comparisons included correlation with usability severity scores and comparative odds of having given a low, medium, or high safety severity score.

Results

A total of 21 unique usability issues corresponding to 58 heuristic violations were identified across the seven tasks. →Fig. 2 summarizes the findings of Phase 1. The bars on the left of the figure depict the total number of heuristic

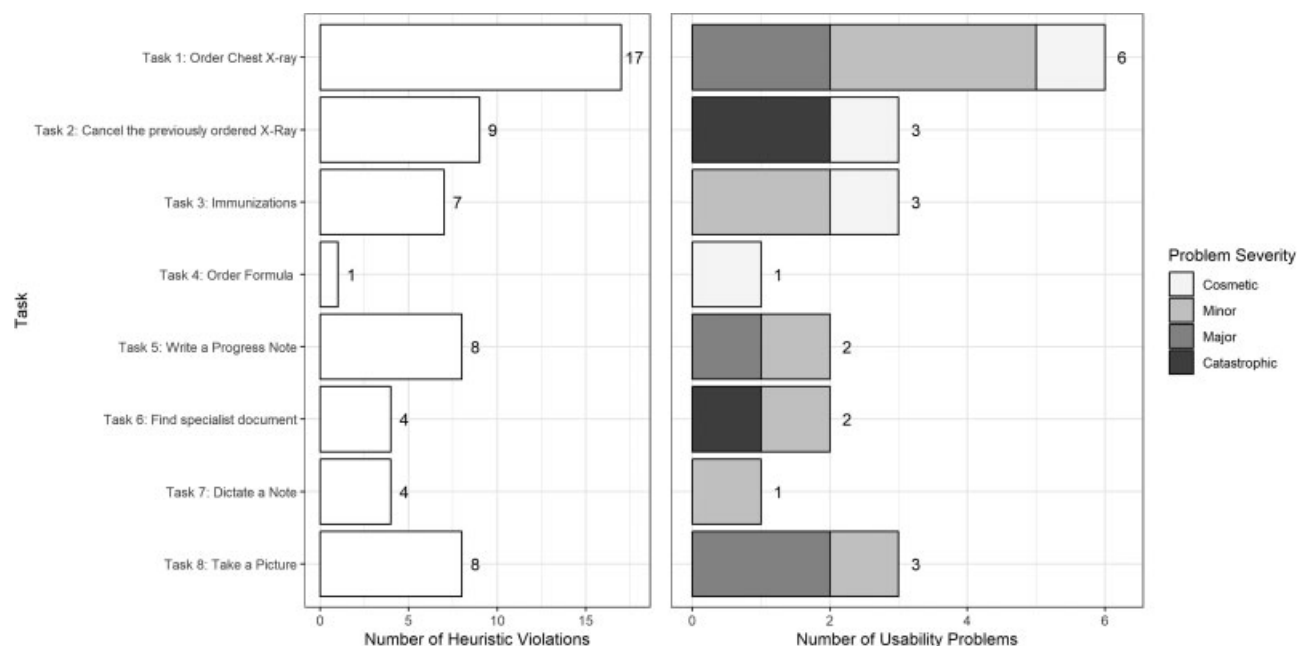


Fig. 2 The summary of heuristic violations and usability problems identified by task. Shading for usability issues on the right depicts various frequencies of problem severity.

violations for each task, whereas the bars on the right side of the figure show the usability issues based on the corresponding heuristic violations. The shading of the bars on the right also depicts the severity score and frequency of that score for each usability problem (see [Fig. 2](#)). Task 1, for example, had 17 heuristic violations resulting from six usability issues; some of the usability issues were associated with multiple heuristic violations. The shading on the right for task 1 indicates that severity ratings revealed one usability issue rated as cosmetic, three usability issues rated as minor, and two usability issues rated as major. None of the tasks was free of heuristic violations. Dictating a note resulted in the least number of usability issues at 1. Order of a chest X-ray resulted in the highest number of heuristic violations at 17. In terms of severity ratings, canceling an order and locating a specialist note were the only tasks that received heuristic severity scores of 4 (catastrophic). Ordering a chest X-ray, starting a progress note, and capturing a picture all received at least one heuristic violation severity score of 3 (major).

[Table 3](#) shows a sample heuristic evaluation by summarizing a portion of the analysis result for task 2, canceling an order for a chest X-ray. It shows heuristic violations corresponding to a single usability issue with usability severity scores and explanations as well as safety severity scores with comments/explanations. When ordering and canceling using the application, if a user submitted a new order or canceled an order without signing, it was possible for the user to leave the screen and the app without any prompt from the app that orders were left unsigned. These orders would in effect never post to the EHR. This problem is a violation of the heuristics of visibility of system state, informative feedback, and clear closure because the user assumes that the order has been submitted but is not

presented with any clear confirmation that submission of the order has been completed. Under the circumstance of placing an urgent order, if that order never posts to the EHR and if the provider is unaware of this, the result could be a delay in care and possible harm to the patient. This usability issue was given a heuristic severity scale of 4 and an independent safety scale rating of 3 to 3.75 ([Table 3](#)), indicating a catastrophic usability problem on the heuristic severity scale that could have major safety consequences.

Phase 2 involved applying safety severity scale scores to the corresponding 28 usability issues. From a safety severity standpoint, some differences were noted in the scoring tendencies between the physician clinical informatics specialists and the clinical safety professionals. Cohen's κ was run on two occasions, first to determine agreement between the two clinical informatics specialists' judgments and second to determine agreement between the two safety officers' judgments. There was moderate agreement between the two safety officers' judgments, $\kappa = 0.580$ (95% confidence interval [CI]: 0.351–0.809), $p < 0.001$. However, the clinical informatics specialists' judgments were not in statistically significant agreement, $\kappa = 0.216$ (95% CI: 0.005–0.427), $p = 0.164$. Minimal agreement was found when all four raters were assessed together, $\kappa = 0.336$, $p < 0.001$. Overall, clinical safety professionals tended toward the extreme ratings of either minor or catastrophic, whereas clinical informatics specialists were more conservative, leaning toward scores of minor or moderate ([Fig. 3](#)). Clinical informatics specialists were about half as likely to have given a low safety severity score (<1.5) as safety professionals (odds ratio [OR]: 0.47 [0.16, 1.40]; $p = 0.18$), although this difference was not statistically significant. Clinical informatics specialists were approximately eight times as likely to have given a medium score (1.5–2.4) as

Table 3 Example of a usability problem by task, heuristics violated, and severity rating

Task	Task 2: cancel the previously ordered X-ray		
Heuristics violated	Visibility of system state	Informative feedback	Clear closure
Description of heuristic violation	Once you click on “discontinue” on the order, the next screen shows the order displayed with strikethrough the order text. You can leave the Orders screen <i>and</i> the application without it prompting for signature. It <i>will</i> prompt you only if you go to another patient. If you do not sign it, it will not go through	Poor feedback. The order appears to be canceled (strikethrough). But if you don’t sign, the cancellation is not finalized.	The only feedback you get is the Sign button and the order disappears. No feedback closure.
Usability issue	When canceling an order, there is no confirmation of cancellation		
Nielsen’s principles	Frequency: common; impact: difficult to overcome; persistence: always		
Severity rater type	Usability (physician rater and human factors scientist consensus score)	Safety (maximum among clinical informaticists)	Safety (maximum among clinical safety professionals)
Severity score	4	3	3.75
Explanation of rating: subjective anecdotal comments by evaluators	It will let you exit the orders screen or the App. It will not prompt you that you have orders waiting to be signed.	Patients could be allergic to meds, or the wrong patient could receive a med and the doctor could think it was canceled. Based on past experience, such things have been fatal.	That is scary. Could be a medication and the nurse would not know. I don’t like that. That is bad. Could be meds, procedures, discharge even. Nothing to prompt that you need to sign or not complete. Example: need to take the patient off pressors, but the order never canceled, leading to patient morbidity

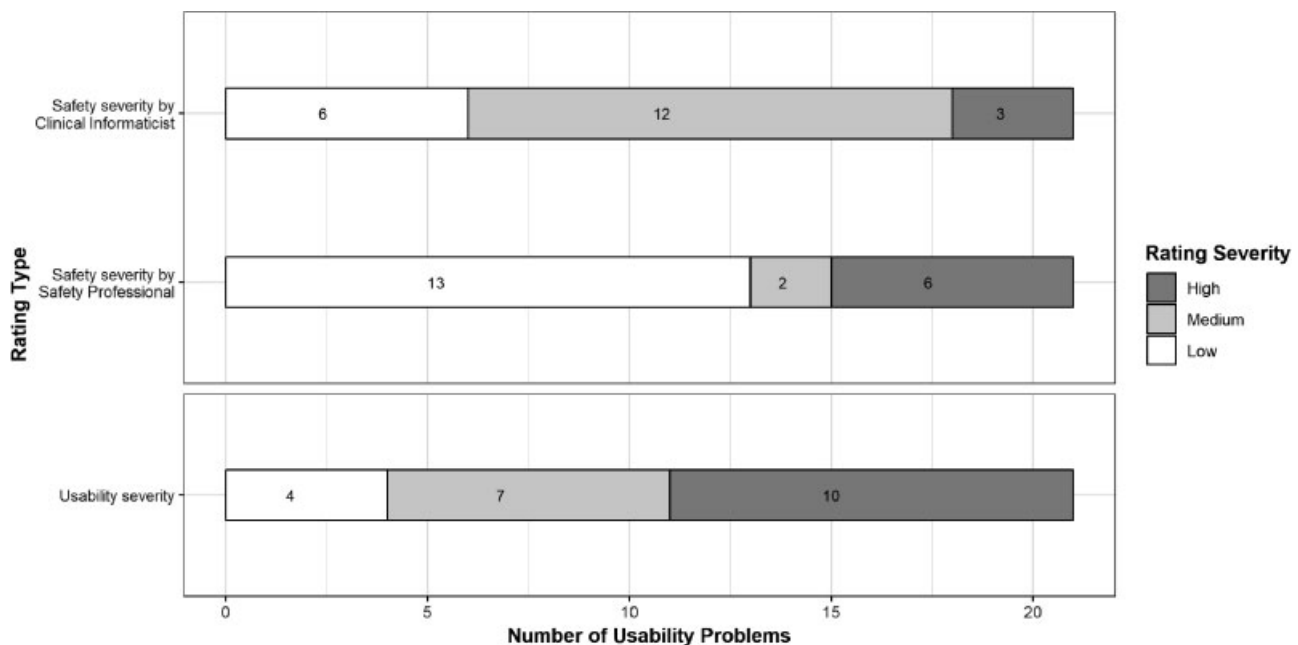


Fig. 3 Side-by-side comparison of ratings between safety (top) and usability (bottom). The safety severity ratings are further broken down to demonstrate the difference between clinical informaticists and safety professionals.

safety professionals (OR: 8.41 [1.65, 42.76]; $p < 0.01$). Finally, the clinical informatics specialists were more than 75% less likely to have given a high score (> 2.5) as safety professionals (OR: 0.23 [0.04, 1.23]; $p = 0.07$), although this difference was only significant at the 0.1 α level.

The primary aim of this study was to investigate a possible relationship between heuristic severity scale ratings and safety scale. The results of this study demonstrate a positive correlation between the heuristic severity scale score and the NIST safety severity scale in that as heuristic severity increased, safety risk also increased. Based on linear models equating the maximum safety severity score for each rater type to the usability score across the 28 usability issues, 49% of the variation of the safety risk score given by clinical safety professionals ($r = 0.70$; $F = 11.04$; $p < 0.01$) and 42% of the variation of the safety risk given by clinical informatics specialists ($r = 0.65$; $F = 19.18$; $p < 0.01$) is explained by the usability severity score of the problem outlined by the heuristic analysis/violation (—Fig. 4).

Discussion

The goal of this study was to try to elucidate whether there was a relationship between EHR design, specifically usability, and possible safety risk for patients. This study shows a positive correlation between usability severity scoring and safety severity scoring across the studied mobile EHR tasks. These findings reinforce the understanding that EHR design may pose potential safety risks for patients and show that careful analysis of EHR structure and workflow may be

necessary to identify failures in EHR design that need to be addressed by designers/engineers. A recent systematic review of usability evaluations of EHRs by Ellsworth et al³² highlights a continuing lack of research proposing specific tools from usability evaluations that can be used effectively for EHR development. Our study suggests that EHR designers could leverage heuristic evaluation by a trained interdisciplinary team as a development tool in designing any portion of the EHR, especially for workflows tied directly to patient care. To refine this process so that it can be used as an effective evaluation process, future investigation is warranted to see if a threshold on the heuristic severity scale can be established, where any heuristic violation severity scale values above the established threshold could be leveraged to prioritize reengineering of existing EHR tasks or for evaluating new EHR software designs prior to deployment. This process could prove to be an effective tool and aligns with the goal of trying to establish higher reliability systems in health care. This approach could offer a relatively inexpensive method for improving interface usability while potentially improving patient safety by avoiding compromised designs. As the EHR application in this study is in ongoing development, it should be noted that some design changes have been made based on some of these findings.

Results from this study also highlighted the importance of using an interdisciplinary team of evaluators in that safety severity scoring varied somewhat by the role of the evaluator. It was noted, however, that this variation was limited by the small number of collaborating experts. Safety officers, for example, tended to rate severity at the extreme of “0”, no

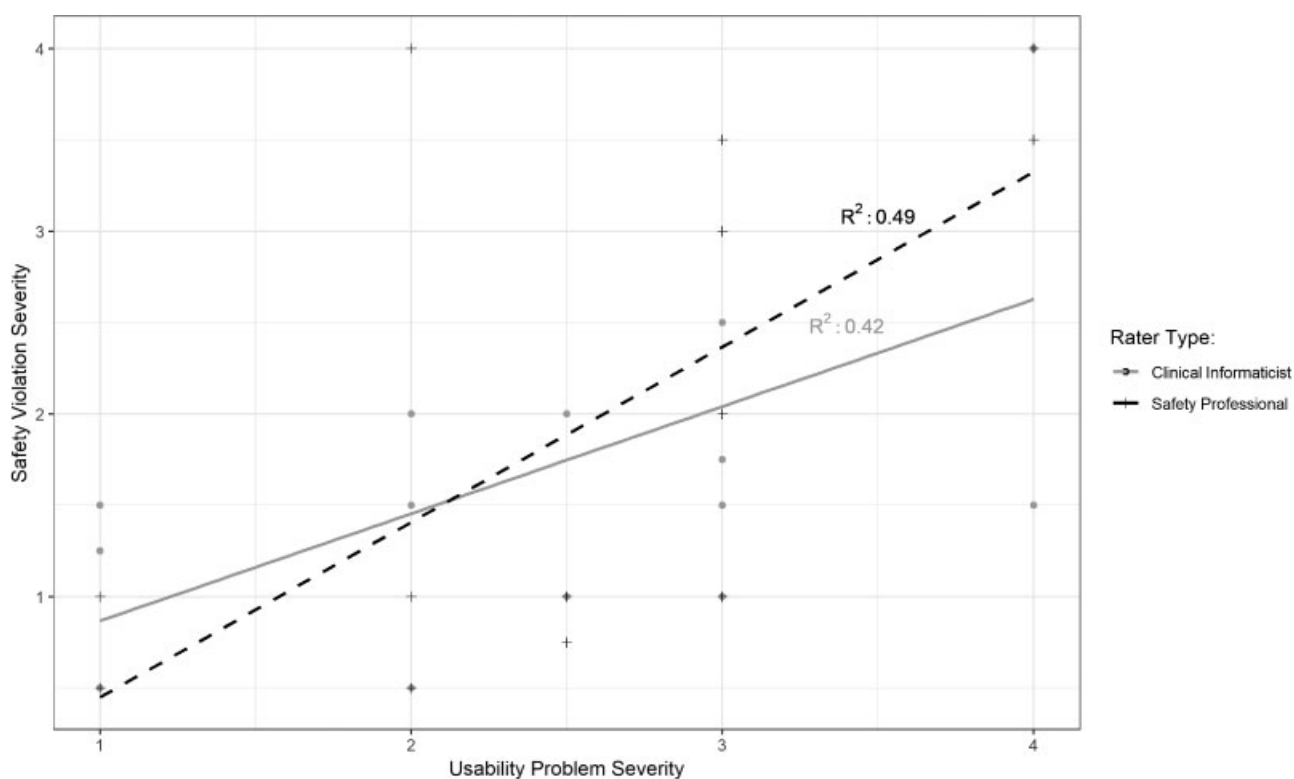


Fig. 4 Positive correlation between usability problem severity ratings and safety severity ratings for both clinical informaticists and safety professionals.

problem, or “4,” catastrophic with potential for patient mortality. Clinicians, on the other hand, tended to rate safety severity at “1,” minor, or “2,” moderate. This discrepancy was reflected in the safety officers’ comments during the evaluation that indicated the safety officers used prior safety event knowledge to guide their ratings. Their comments anecdotally referenced previous safety events tied to the EHR that were similar to analyzed tasks and usability problems. For example, during the safety analysis session, one safety officer commented, “I have seen events where an order was not placed or not canceled due to a problem with the EHR.” The involvement of certified safety officers in phase 2 of the study brought a novel and valuable perspective to the potential safety impact of usability. Due to their training and broad experience in investigating safety events within health care, their point of view focuses more on attempting to understand the safety implications of the noted usability problems from a global system perspective. Although safety officers do not routinely interact with the EHR in direct patient care and therefore would not routinely be targeted as end users for usability evaluations, safety officers may see the downstream effects of poor EHR design. The clinical informatics physicians who performed the safety scales in our study did not have any formal health care safety training nor did they participate frequently in safety event investigation. Their point of view therefore relied on their clinical experience and informatics expertise and was reflected in their scoring. This difference in perspectives particularly highlights the importance of using an interdisciplinary group of experts to perform each area of the evaluation and further understanding of how the evaluator’s role might affect severity scoring may be valuable in future research.

There are several limitations to this study. It should be emphasized that this study does not provide the full picture of the potential risk of EHR interfaces, causing harm to patients. Usability testing still needs to be conducted as a complementary method to uncover all aspects of design that could contribute to inadvertent user errors. Our study had a relatively small number of expert collaborators, and relatively few tasks were evaluated. The lack of medication orders within the app is important as well due to the relatively high number of safety events that are a result of computerized medication order entry. Further investigation is needed to validate these results with higher numbers of expert reviewers. Future studies may consist of more detailed assessment across different EHR vendors and platforms (desktop vs. mobile) with more detailed and broadened task lists, and tasks undertaken by other roles within health care, such as nursing care or pharmacy. Future investigations may also extend our understanding of potential usability issues by correlating safety ratings with observed errors.

Finally, this study raises the question of whether there should be further official recommendations or regulations from federal regulators that require industry to incorporate more rigorous processes for evaluation and assessments of EHR design. The burden of assessing these designs must be balanced by the risk of stifling ongoing innovation for more efficient and user-friendly interfaces.

Conclusion

From this study, we can draw three conclusions. First, there appears to be a positive correlation between EHR usability and its potential for affecting safety in patient care. Second, it appears that based on this relationship, heuristic evaluation and usability violation scale severity scores could potentially ameliorate potential safety risks during EHR design processes prior to application deployment as well as aid in prioritizing already deployed architectures for reengineering. Finally, we conclude that the use of a diverse interdisciplinary team for investigating usability involving experts in human factors, medical practice, and safety is an important approach to establishing higher reliability systems of evaluating EHR function and usability and its relationship with patient safety—a top priority in health care. This approach could have great potential for reducing downstream patient safety risk while refining an interface that would ultimately be more user-friendly. In addition, if applied early in the design and development of digital solutions, this approach could lower the costs by reducing the need for software reengineering downstream patient safety risk while refining an interface that would ultimately be more user-friendly. In addition, if applied early in the design and development of digital solutions, this approach could lower the costs by reducing the need for software reengineering downstream.

Multiple Choice Questions

1. This study suggests that this new process of safeheuristics is optimally performed by:
 - (a) Safety experts.
 - (b) Frontline users.
 - (c) Legal experts.
 - (d) Human factors/ergonomics experts.
 - (e) Inter-disciplinary teams of experts.

Correct Answer: The correct answer is option e.

2. The study results suggest that the safety ratings and usability heuristics have a:
 - (a) Small positive correlation.
 - (b) Moderate positive correlation.
 - (c) Small negative correlation.
 - (d) Moderate negative correlation.
 - (e) No significant correlation.

Correct Answer: The correct answer is option b.

3. This study suggests that usability heuristics account for ___ percentage of the safety scores:
 - (a) < 25%
 - (b) 25–39%
 - (c) 40–50%
 - (d) 51–65%
 - (e) > 65%

Correct Answer: The correct answer is option c.

Conflict of Interest

None declared.

Acknowledgments

The authors would like to acknowledge Stephen Klem, MD, Robert Sturdevant, MD, Rachael Dameron, MBA, MBHS, RRT-NPS, CPPS, and Audrey Kennedy, PharmD, BCPS, CPPS, for their expertise and assistance.

References

- Adler-Milstein J, DesRoches CM, Kralovec P, et al. Electronic health record adoption in US hospitals: progress continues, but challenges persist. *Health Aff (Millwood)* 2015;34(12):2174–2180
- Blumenthal D, Tavenner M. The “meaningful use” regulation for electronic health records. *N Engl J Med* 2010;363(06):501–504
- Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med* 2006;144(10):742–752
- Kruse CS, Kristof C, Jones B, Mitchell E, Martinez A. Barriers to electronic health record adoption: a systematic literature review. *J Med Syst* 2016;40(12):252
- The Office of the National Coordinator for Health Information Technology. Test Procedure for §170.314(g)(3) Safety-enhanced design. Available at: <https://www.healthit.gov/sites/default/files/standards-certification/2014-edition-draft-test-procedures/170-314-g-3-safety-enhanced-design-2014-test-procedures-draft-v-1.0.pdf>
- Couture B, Lillie E, Chang F, et al. Applying user-centered design methods to the development of an mHealth application for use in the hospital setting by patients and care partners. *Appl Clin Inform* 2018;9(02):302–312
- Preece J, Rogers Y, Sharp H. *Interaction Design: Beyond Human-Computer Interaction*. 4th ed. Chichester: Wiley; 2015
- Sorbello A, Ripple A, Tønning J, et al. Harnessing scientific literature reports for pharmacovigilance. Prototype software analytical tool development and usability testing. *Appl Clin Inform* 2017;8(01):291–305
- Tarrell A, Grabenbauer L, McClay J, et al. Toward improved heuristic evaluation of EHRs. *Health Syst (Basingstoke)* 2015;4:138–150
- Nielsen J, Molich R. Heuristic evaluation of user interfaces. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Seattle, Washington, United States, April 1–0, 1990, 249–256. Available at: <http://dl.acm.org/citation.cfm?id=97281>. Accessed February 10, 2017
- Hertzum M, Jacobsen NE. The evaluator effect: a chilling fact about usability evaluation methods. *Int J Hum Comput Interact* 2003;15:183–204
- Jeffries R, Desurvire H. Usability testing vs. heuristic evaluation. *ACM SIGCHI Bull* 1992;24:39–41
- Molich R, Nielsen J. Improving a human-computer dialogue. *Commun ACM* 1990;33:338–348
- Zhang J, Patel VL, Johnson TR, et al. Evaluating and predicting patient safety for medical devices with integral information technology. In: Henriksen K, Battles JB, Marks ES, et al., ed. *Advances in Patient Safety: From Research to Implementation (Volume 2: Concepts and Methodology)*. Rockville, MD: Agency for Healthcare Research and Quality; 2005
- Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. *J Biomed Inform* 2003;36(1-2):23–30
- Cooper JB, Newbower RS, Long CD, McPeck B. Preventable anesthesia mishaps: a study of human factors. *Anesthesiology* 1978;49(06):399–406
- Graham MJ, Kubose TK, Jordan D, Zhang J, Johnson TR, Patel VL. Heuristic evaluation of infusion pumps: implications for patient safety in Intensive Care Units. *Int J Med Inform* 2004;73(11-12):771–779
- Allen M, Currie LM, Bakken S, Patel VL, Cimino JJ. Heuristic evaluation of paper-based Web pages: a simplified inspection usability methodology. *J Biomed Inform* 2006;39(04):412–423
- Tang Z, Johnson TR, Tindall RD, Zhang J. Applying heuristic evaluation to improve the usability of a telemedicine system. *Telemed J E Health* 2006;12(01):24–34
- Bates DW, Leape LL, Cullen DJ, et al. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. *JAMA* 1998;280(15):1311–1316
- Kaushal R, Shojania KG, Bates DW. Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. *Arch Intern Med* 2003;163(12):1409–1416
- 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
- Walker JM, Carayon P, Leveson N, et al. EHR safety: the way forward to safe and effective systems. *J Am Med Inform Assoc* 2008;15(03):272–277
- Han YY, Carcillo JA, Venkataraman ST, et al. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. *Pediatrics* 2005;116(06):1506–1512
- Graber ML, Byrne C, Johnston D. The impact of electronic health records on diagnosis. *Diagnosis (Berl)* 2017;4(04):211–223
- Sittig DF, Wright A, Ash J, et al. New unintended adverse consequences of electronic health records. *IMIA Yearb* 2016;(01):7–12
- Sittig DF, Ash JS, Singh H. Overview of safer guides. *SAFER Electron Health Rec Saf Assur Factors EHR Resil* 2015;20:153–156
- Sittig DF, Singh H. Toward more proactive approaches to safety in the electronic health record era. *Jt Comm J Qual Patient Saf* 2017;43(10):540–547
- Lowry S, Quinn M, Ramaiah M, et al. (NISTIR 7804) Technical Evaluation, Testing and Validation of the Usability of Electronic Health Records. NIST Interagency/Internal Report (NISTIR) - 7804. 2012. Available at: <https://www.nist.gov/node/592206>. Accessed February 10, 2017
- Borycki E, Kushniruk A, Carvalho C. A methodology for validating safety heuristics using clinical simulations: identifying and preventing possible technology-induced errors related to using health information systems. *Comput Math Methods Med* 2013. Doi: 10.1155/2013/526419
- Nielsen J. Severity Ratings for Usability Problems. Nielsen Norman Group. 1995. Available at: <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>. Accessed April 2, 2018
- Ellsworth MA, Dziadzko M, O'Horo JC, Farrell AM, Zhang J, Herasevich V. An appraisal of published usability evaluations of electronic health records via systematic review. *J Am Med Inform Assoc* 2017;24(01):218–226