ISSN: 2321-8819 (Online) 2348-7186 (Print) Impact Factor: 1.498 Vol. 6, Issue 2, February, 2018

Usability Assessment of an EHR's Clinical Notes Interface from Attendings and Residents Vision: An Exploratory Study

Farah Sattar Choudhry[1], Muhammad Yousaf[2] and Muhammad Usman Zafar[3]
[1]Allama Iqbal Medical College, Lahore, Pakistan
Email: farahsattar21@gmail.com
PMDC# 85297-P
[2]Services Institute of Medical Sciences, Lahore, Pakistan
Email: yousaf164@gmail.com
PMDC# 85881-P
[3]Dali University, China
Email: coolmanusman@gmail.com
PMDC# 78800-P

Abstract

Background: Usability gaps between current and future improved Electronic Health Record (EHR) system designs exist due to insufficient incorporation of User-Centered Design (UCD) principles during System Development Life Cycle (SDLC).

Objectives: To evaluate the usability aspects of a commonly used EHR system specific to clinical notes usage from attendings' and residents' standpoints by analyzing objective measures of users' performance and their subjective perceptions employing mixed methods approach.

Methods: Usability of a commercial, inpatient EHR clinical notes documentation interface was analyzed from standpoints of two provider groups employing two standardized patient cases. Both objective and subjective data were collected from attending (n=6) and resident physicians (n=8) through usability testing employing a mixed method approach.

Results: The study results suggested that (i) EHR usability and desirability is influenced by user characteristics, (ii) workloads associated with H&P and progress notes writing are perceived differently between two groups, (iii) repeated task performance improves user efficiency and (iv) user performance is correlated to their subjective system assessments.

Conclusion: Understanding usability of clinical documentation interface from perspectives of two different user groups, provides interface designers with an opportunity to develop an EHR system centered on UCD principles.

Keywords: Electronic Health Record (EHR); User-Computer Interface; Documentation

Introduction

While Electronic Health Record (EHR) systems have been widely adopted with the ultimate goal of improved health care delivery (1), substantial gaps exist between the current state of EHRs and their potential usefulness (2). Poor EHR usability appears to be a major factor for this discrepancy (2). To facilitate optimal end product usability, it is critical to understand end users' "usage behavior", considered a core feature of a User- Centered Design (UCD) approach (6,92). The UCD philosophy is that "the final product should suit the users, rather than making the users suit the product" (70). According to the International Organization for Standardization (ISO)-framework used in this research study, usability is defined as the, "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (93). Similarly, in EHR design, user involvement throughout the System Development Life Cycle (SDLC) can facilitate the development of systems that are easy to learn and remember, efficient, minimize errors and improve user satisfaction (94), which could improve EHR adoption and better patient outcomes (9)

Despite the critical role of the Human Computer Interaction (HCI) in the SDLC process (66), it is often neglected during EHR interface design. Usability studies on EHRs' clinical decision support system and user interfaces for medical equipment have been done in the past (21,22), but there are not many studies focusing on clinical notes documentation within an EHR interface (24-27), with only few studies done on usability evaluation and prototyping of clinical notes user interfaces in the medical domain (24- 29). Similarly, usability of a system could vary with vendor types and user profiles (e.g., clinical experience, EHR training, age, gender, technology skills). However, few research studies incorporate usability comparisons from diverse user perspectives (e.g., expert users vs. novice users; physician vs. patients; users vs. usability experts) (34).

Usability testing is accepted as the most effective usability methodology with greatest strategic impact (45). It is an "activity that focuses on observing users working with a product and

performing tasks that are real and meaningful to them" (66). The purpose of this study is to quantify EHR usability around inpatient notes usage focusing on the clinical note documentation and clinical note viewing interface, an area that poses tremendous challenges to physicians and other clinicians working under time limitations (25). Both objective and subjective data on users' task performance were collected from two user groups (i.e., attendings and residents) and analyzed via usability metrics as defined by ISO (i.e., effectiveness, efficiency and satisfaction) (33). Supplementary data were also analyzed for subjective workload using the NASA-TLX instrument (26) and system desirability with Product Reaction Cards (PRC) (37). The insight gained through this research provides an opportunity to better understand EHR usability around clinical documentation from the standpoints of two provider groups and identify usability gaps to benchmark future EHR design.

Methods

This research study evaluated the usability of an enterprise EHR (Epic Systems Corporation) system at Fairview Health Services, University of Minnesota Medical Center (UMMC). The study specifically focused on clinical documentation tasks (e.g., H&P and progress note-writing). Scenario-based usability testing was conducted on two

high fidelity simulated test patient charts (22) in an Epic test environment replicating the real work environment, both in design and functionality. Testing was done at the usability laboratory.

Study Sample

Physician participants (n=14) were from two user groups: attendings (n=6) and residents, excluding interns (n=8). Participants were in all cases either trained in Internal medicine or Family medicine with past and/or current inpatient experience with the Epic Fairview EHR. Detailed user characteristics categorized are summarized in Table 1.

Table 1-Users characteristics

	Age (M/F)	Clinical Exp. (Yrs.)	Technology Exp.	Epic Exp. (Yrs.)	Epic proficiency
A	31 (F)	=5	Somewhat	5-10	Average
A	43 (F)	>10	Less	5-10	Average
A	xx (F)	=5	Somewhat	5-10	Average
A	43 (F)	>10	Very	>10	Proficient
A	36 (M)	=5	Somewhat	<5	Proficient
A	39 (M)	=5	Somewhat	5-10	Average
R	30 (F)	=5	Somewhat	<5	Average
R	xx (F)	=5	Somewhat	<5	Average
R	xx (M)	=5	Somewhat	<5	Average
R	30 (F)	=5	Somewhat	<5	Average
R	29 (M)	=5	Somewhat	<5	Proficient
R	26 (M)	=5	Very	<5	Proficient
R	29 (M)	=5	Somewhat	<5	Proficient
R	29 (F)	=5	Somewhat	<5	Average

A=Attendings; R=Residents; Clinical Exp.: Clinical Experience (Residency training and later); Epic Exp.: Total years of experience using Epic

Participation was voluntary and participants received $50/hour. Each session was 2.5-3 hours long and each physician was at least 24 hours off night call on his or her day of data collection. The study protocol was approved by the Institutional Board Review.

Data Collection and Analysis

Two simulated, high fidelity test patient charts with rich, realistic clinical data were created in an Epic test environment to provide scenario-based EHR usability testing (38). Patients were built from representative cases after extensive discussion among five experts: the lead EHR physician trainer (MS) and four physicians in formaticists (RR, TA, GMM & EA). Patient cases with similar complexities were selected using a Charlson weighted comorbidity index and number of prior admissions, clinic visits, and clinical notes. In both clinical scenarios, patients with a history of hronic Obstructive Pulmonary Disease and Congestive Heart Failure presented in the emergency

department with sudden onset shortness of breath. Each participant was assigned two patient cases in a random order employing an online randomization tool (43). A Randomized blocked design approach was used to create balanced distribution of test patients across two groups. Each participant performed the same tasks of entering a H&P and a day 1 progress note, on each test patient's chart.

Raw data was extracted employing Tobii studio version 3.4.5 and was evaluated in three ways: (a) user satisfaction, via the System Usability Scale (SUS) questionnaire (100,101) (b) efficiency, via time on tasks, key presses, & mouse clicks and (c) effectiveness, via note quality using the Physician Documentation Quality Instrument-9 (PDQI-9) (23) and overall Gestalt judgment (43). Data from each user group was also analyzed for subjective workload index using the NASA-TLX questionnaire (32) and system desirability via Product Reaction Cards (PRC) listing 118 words (33). All participants were asked to circle their top 5 choices, which were later compiled as a word cloud and Venn diagram to visualize total and unique word selection by each user group.

Note quality assessment was performed by two co-authors/physicians (RR and TA), using standardized metrics as previously reported with the Physician Documentation Quality Instrument-9 (PDQI-9) (33) and overall Gestalt judgment (12). Pretesting of these instruments for note quality assessment was conducted on a set of unrelated notes to ensure that both reviewers shared a common understanding of item scoring. Once consensus was achieved, both evaluators reviewed and assessed approximately 14% of notes (8 of 56 notes). The consistency in quality assessment was checked by inter-coder agreement with final mean agreement for PDQI-9 of 81% (kappa=0.69) and Gestalt scoring of 87.5% (kappa=0.71). We report summative statistics using SAS enterprise guide 5.1 and StatPlus LE 6.0.3 (a statistical software plugin for Macintosh), with means and standard deviation (sd).

Results

While not statistically different, user satisfaction with respect to overall usability of clinical note documentation was perceived worse by attendings (mean SUS = 60.8 ± 15.6 (i.e., marginal usability)), compared to residents (mean SUS = 73.4 ±13.5, (i.e., acceptable usability)), despite longer average Epic experience among attendings (≥ 5 years, n=5/6) compared to residents (< 5 years, n=8/8). The SUS and their interpretation

(101) are illustrated in Fig.1.

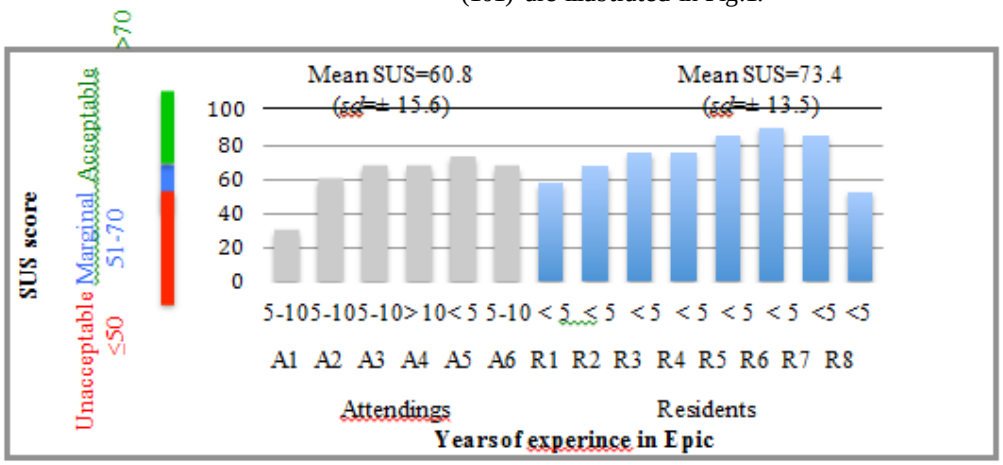

Figure 1-SUS based on users' characteristics

Efficiency was quantified based on time on task, key presses, and mouse clicks. H&P writing was more time-intensive than progress notes for both attendings (26.2 ± 9.7 vs. 14.0 ± 6.4 minutes) and residents (24.2 ± 7.7 vs. 12.3 ± 4.5 minutes). Residents took slightly less time than attendings writing both H&P (24.2 ± 7.7 vs. 26.2 ± 9.7 minutes) and progress notes (12.3 ± 4.5 vs. 14.0 ± 6.4 minutes). Time on task decreased from the 1st to 2nd patient, except for progress note-writing among residents (Fig. 2).

ISSN: 2321-8819 (Online) 2348-7186 (Print) Impact Factor: 1.498 Vol. 6, Issue 2, February, 2018

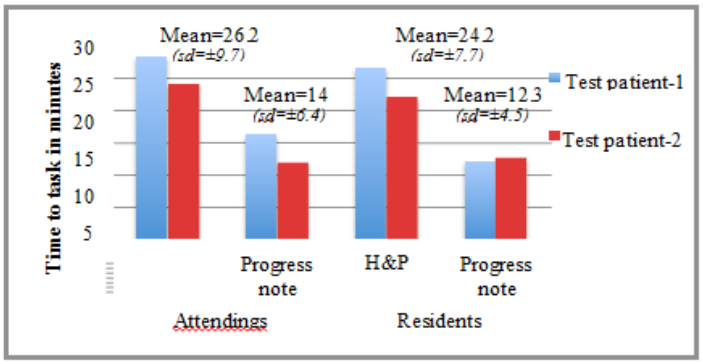

Figure 2-Time to tasks comparison between two user groups

More key presses (KP) and mouse clicks (MC) were observed with H&P as compared to progress note-writing for both attendings (KP=2,644 ± 1535 vs. 1,433 ± 682, MC=201 ± 83 vs. 126 ± 60) and residents (KP=3,468 ± 1,199 vs. 1,758 ± 689

MC=214 ± 82 vs. 112 ± 46) with residents generally performing more key presses and mouse clicks compared to attendings with exception of progress notes where attendings had more mouse clicks. The number of key presses and mouse clicks decreased from the 1st to 2nd patient, except for number of mouse clicks by residents during progress note-writing (Fig. 4.3, 4.4).

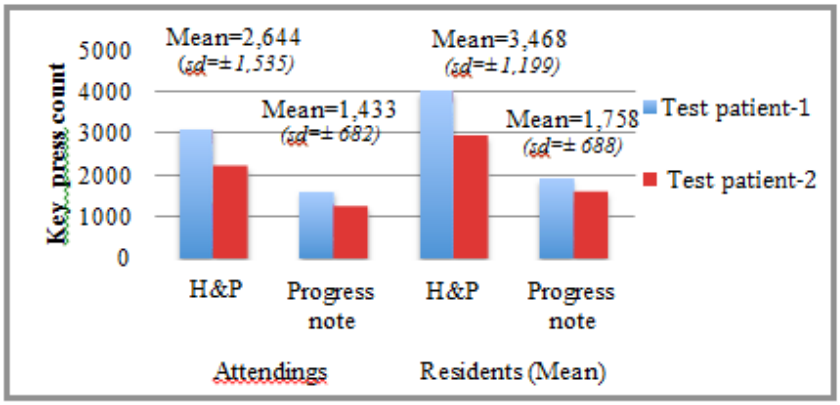

Figure 3-Number of key presses comparison between two user groups

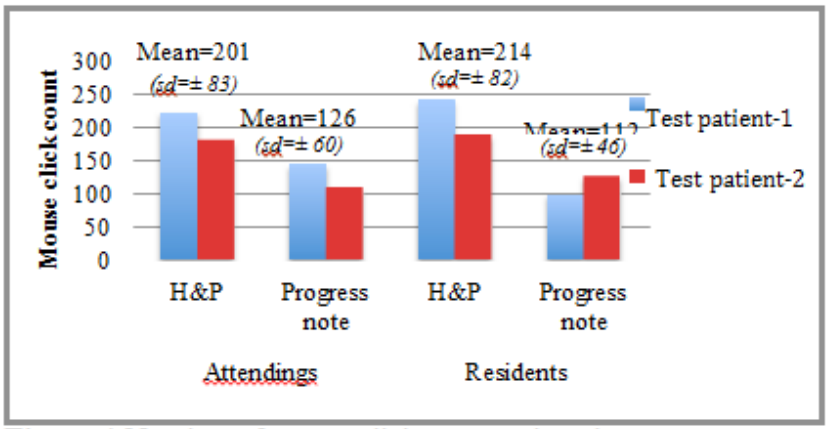

Figure 4-Number of mouse clicks comparison between two user groups

Effectiveness, as measured through PDQI-9 scores on note quality showed no quality differences between H&P and progress notes by attendings (34.9 ± 3.8 vs. 34.8 ± 4.8), though resident progress notes were slightly higher quality than H&P notes (35.5 ±

6.3 vs. 33.8 ± 4.0). Attendings' H&P notes (34.9 ± 3.8 vs. 33.8 ± 4.0) and residents' progress notes (35.5 ± 6.3 vs. 34.8 ± 4.8), showed only minimal quality differences. No noticeable differences in note quality between attending and residents were detected through Gestalt scoring both for H&P (3.7 ± 0.7 vs. 3.8 ± 0.8) and progress notes (3.9 ±

0.9 vs. 4.0 ± 1.0). PDQI-9 scores increased from the 1st to 2nd patient, except for residents' progress notes (Fig. 5).

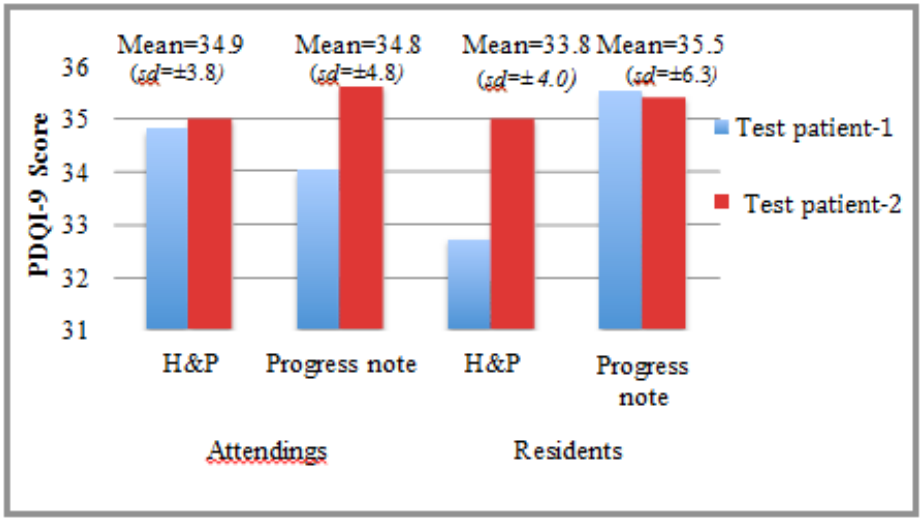

Figure 5-PDQI-9 scores comparison between two user groups

The NASA-RTLX questionnaire revealed that H&P note-writing had higher overall workload (OW) than progress note-writing among both attendings (27.8 ± 11.4 vs. 27.2 ± 16.0) and residents (33.6 ± 16.7 vs. 22.5 ± 10.2). Residents also had considerably higher subjective OW for H&P note-writing (33.6 ± 16.7 vs. 27.8 ± 11.4), while attendings had higher subjective OW for progress note-writing (27.2 ± 16.0 vs. 22.5 ± 10.2). There was no effect of patient order on perceived workload (Fig. 6).

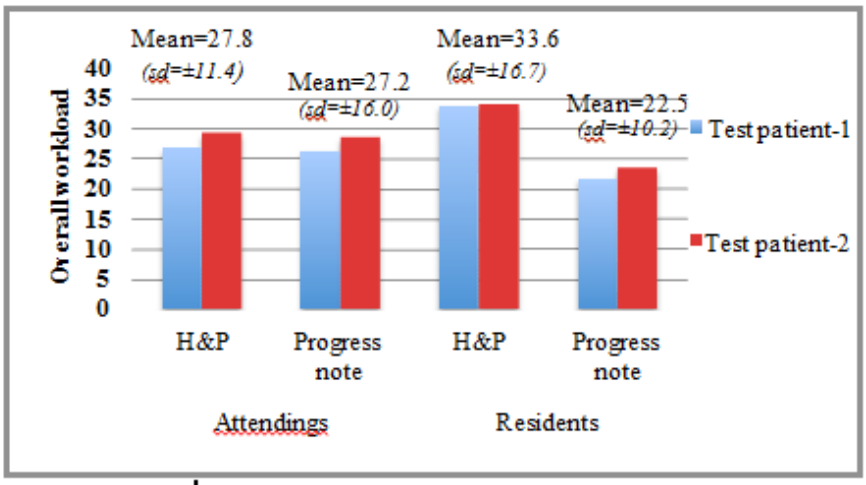

Figure 6-NASA-RTLX scores comparison between two user groups

Pearson correlation coefficient calculations were performed on the data after visually inspecting the distributions for normality. The results showed some correlation between metrics: NASA & SUS (-0.79 vs. -0.55 i.e., strong negative), NASA & Gestalt (-0.27 vs. -0.26 i.e., fair degree of negative), PDQI-9 & Gestalt (0.82 vs. 0.70 i.e., strong positive), Gestalt & SUS (0.39 vs. 0.30 i.e., fair degree of positive). Other metrics showed weak or no correlation: PDQI-9 & SUS (0.13 vs. 0.14), time on task & PDQI-9 (-0.18 vs. 0.10) and time on task & Gestalt (0.23 vs. 0.20). System desirability analysis compared the proportion of positive vs. negative terms from a comprehensive list of 118 words (97). A higher percentage of positive as compared to negative words were selected both by attendings (63% vs. 37%) and residents (73% vs. 28%). Attendings selected a higher percentage of negative words (37% vs. 28%) while residents selected a higher percentage of positive words (73% vs. 63%) words respectively as depicted in the word cloud images (Fig 7). Similar results were seen for unique word selection as shown in the Venn diagram.

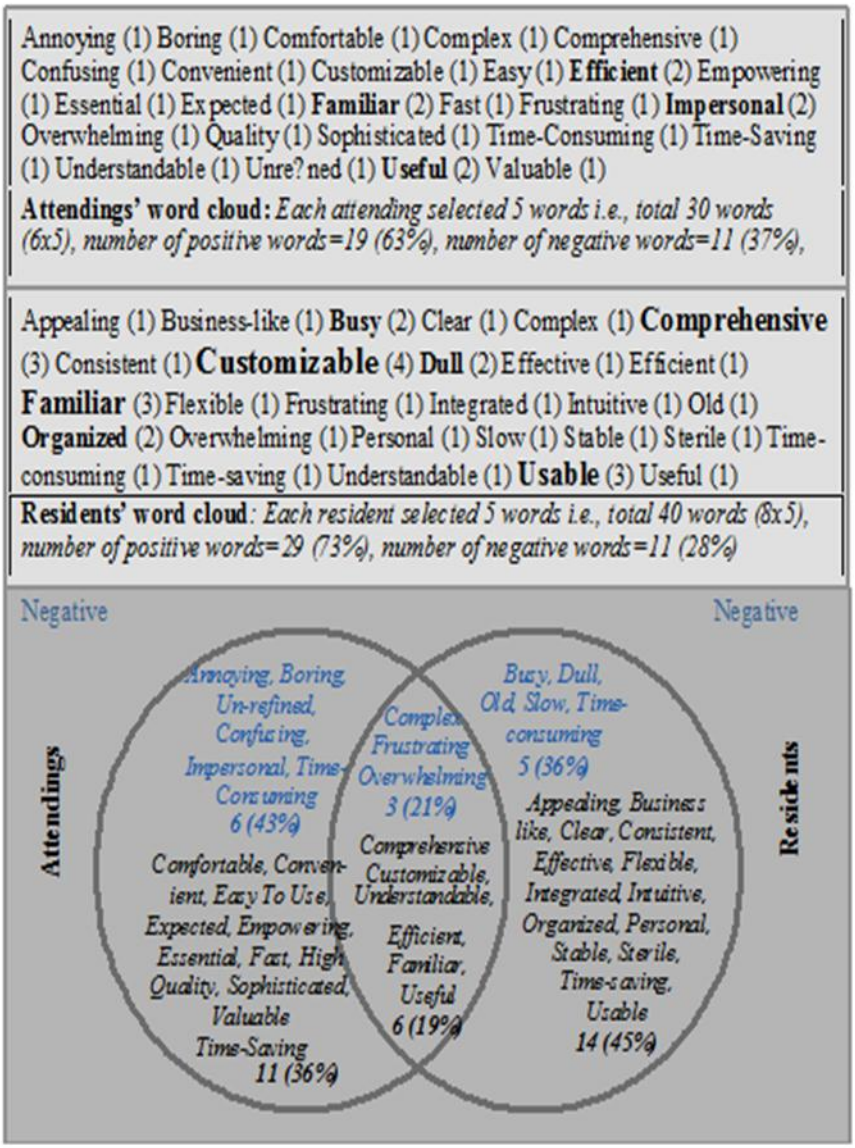

Figure 7-Product Reaction Cards selected by attendings and residents

Discussion

This research study is an important initial step towards understanding the usability of EHR clinical notes documentation from attending and resident physician perspectives. EHR usability, as quantified through objective measures of user performance and their subjective perceptions, varied with each group, note type, and repeated tasks. Varying degrees of correlation were also

discovered between variables, suggesting that user performance is related to their subjective system assessments. The insight gained through this research provides an opportunity to better understand EHR clinical documentation usability, identify and address existing usability gaps, and establish benchmarks for future EHRs.

Based on the SUS, residents perceived the system to have "acceptable usability" while attendings perceived the system to have "marginal usability", despite attendings having more experience with Epic than residents. Similarly, system desirability was considered better among residents compared to attendings, with a higher percentage of positive words used to describe the system. Since residents are generally exposed to EHRs early in their medical training and tend to have little exposure to traditional paper charting, this may explain more favorable responses to EHR usability and desirability. Additionally, resident participants, predominantly males, tended to be younger, and rated themselves as having more technical experience, leading to easier technology adoption (Table 1). Thus, user characteristics appear to be a critical factor for EHR usability.

In terms of efficiency, as quantified by time on task, key presses and mouse clicks, attendings and residents both took significantly more effort with H&P compared to progress note-writing. Residents perceived less subjective workload associated with progress notes suggesting that residents were more at ease in writing progress notes. A potential reason for this is the nature of progress note-writing task itself, which is more repetitive and most likely to be influenced by a system's usability (e.g., copying and pasting, auto population, multiple screen panel functionalities.). In comparison, attendings showed less subjective workload with H&P writing suggesting that they are better skilled in writing H&P notes, a cognitively demanding task which involves providing a reason for admission and providing initial patient management direction. Thus, targeted note documentation training of physicians where there is a lack of proficiency (e.g., H&P among residents and progress note in attendings), would be a reasonable approach to consider. No noticeable difference in note quality between attending and residents was detected through Gestalt scoring.

Generally, efficiency improved as users performed the same note-writing tasks on the 2nd patient with the exception of progress note-writing among residents. The plausible explanation of the observed differences may be due to user familiarity with the systemand faster cognitive processing as a result of repeated task performance, as well as specifics around the second patient case. No effect of patient order was observed on perceived workload while there was some indication of improvement in note quality, especially progress note documentation among attendings and H&P writing among residents.

We also discovered that increases in subjective workload (NASA) were associated with decreases in user satisfaction (SUS) and note quality (per Gestalt). Higher satisfaction (SUS) was associated with better quality notes (per Gestalt). We found a strong positive correlation between PDQI-9 & Gestalt, but no correlation was detected between PDQI-9 & SUS or with time on task and note quality for both PDQI-9 & Gestalt.

There are some limitations associated with this study, including a small sample size lacking significant inferential statistical results. Generalizability is limited due to the inclusion of physicians (MDs), with training in either Internal medicine or Family medicine, and testing of inpatient EHR interfaces only. Additionally, the impact of other user characteristics needs to be explored further. There are some limitations associated with usability testing itself, due to individual differences among users, relevance of tasks being tested, and system speed and connectivity. Future studies with larger sample sizes, more diverse groups of users and tasks, and extension beyond inpatient clinical notes are needed. Also, understanding physicians' EHR usage behaviors around clinical note documentation, the goal of our next study, is an important area that needs to be further explored.

Conclusions

We discovered that EHR usability measures of satisfaction, efficiency, and effectiveness vary with users' characteristics, specific note types, and from repeated performance of the same task on consecutive patients' charts. This study provides preliminary, yet essential information on objective measures of user performance and their perceptions of EHR usability around clinical notes usage. These measures can serve as initial guidance to build EHR interfaces grounded on a "User-Centered Design" approach.

References :

1. Saleem JJ, Plew WR, Speir RC, Herout J, Wilck NR, Ryan DM, et al. Understanding barriers and facilitators to the use of Clinical Information Systems for intensive care units and Anesthesia Record Keeping: A rapid

ethnography. Int J Med Inf 2015;84(7):500-511.
2. McMullen C, Ash J, Sittig D, Bunce A, Guappone K, Dykstra R, et al. Rapid Assessment of Clinical Information Systems in the Healthcare Setting. Methods Inf Med 2011;50(4):299-307.
3. Dorfsman ML, Wolfson AB. Direct observation of residents in the emergency department: a structured educational program. Acad Emerg Med 2009;16(4):343-351.
4. Kopp BJ, Erstad BL, Allen ME, Theodorou AA, Priestley G. Medication errors and adverse drug events in an intensive care unit: direct observation approach for detection. Crit Care Med 2006;34(2):415-425.
5. Timestamped field notes. 2015; Available at: http://www.neukadye.com/mobile- applications/timestamped-field-notes/.
6. Srnka KJ, Koeszegi ST. From words to numbers: how to transform qualitative data into meaningful quantitative results. Schmalenbach Business Review 2007;59(1):29-57.
7. Graphical User Interface. Dec. 2015.; Available at: <http://www.merriam- webster.com/dictionary/graphical user interface>.
8. Farri O, Rahman A, Monsen K, Zhang R, Pakhomov S, Pieczkiewicz D, et al. Impact of a prototype visualization tool for new information in EHR clinical documents. Appl Clin Inform 2012;3(4):404-418.
9. Barnum CM, Barnum CM. Usability testing essentials ready, set-- test! Burlington, Mass.; Burlington, MA: Burlington, Mass. : Morgan Kaufmann; 2010.
10. Iso W. 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs). The international organization for standardization 1998;45.
11. Nielsen J. Usability engineering. Boston: Boston : Academic Press; 1993.
12. Nielsen J. Usability 101: Introduction to usability 2003.
13. Baxter K. Understanding your users : a practical guide to user research methods. Second edition.. ed. Waltham, MA: Waltham, MA : Morgan Kaufmann; 2015.
14. Kushniruk AW, Patel VL, Cimino JJ. Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces. Proc AMIA Annu Fall Symp 1997:218-222.
15. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004 2;37(1):56-76.
16. Beynon-Davies P. Ethnography and information systems development: Ethnography of, for and within is development. Information and software technology 1997;39(8):531- 540.
17. Othman R. An applied ethnographic method for evaluating retrieval features. The Electronic Library 2004;22(5):425-432.
18. Kopp BJ, Erstad BL, Allen ME, Theodorou AA, Priestley G. Medication errors and adverse drug events in an intensive care unit: direct observation approach for detection. Crit Care Med 2006;34(2):415-425.
19. Kaipio J, Lääveri T, Hyppönen H, Vainiomäki S, Reponen J, Kushniruk A, et al. Usability problems do not heal by themselves: National survey on physicians' experiences with EHRs in Finland. Int J Med Inf 2017;97:266-281.
20. Zahabi M, Kaber DB, Swangnetr M. Usability and safety in electronic medical records interface design: A review of recent literature and guideline formulation. Hum Factors 2015;57(5):805-834.
21. Georgsson M, Staggers N. An evaluation of patients' experienced usability of a diabetes mHealth system using a multi-method approach. J Biomed Inform 2016;59:115- 129.
22. Khajouei R, Hasman A, Jaspers MWM. Determination of the effectiveness of two methods for usability evaluation using a CPOE medication ordering system. Int J Med Inf 2011 5;80(5):341-350.
23. Walji MF, Kalenderian E, Piotrowski M, Tran D, Kookal KK, Tokede O, et al. Are three methods better than one? A comparative assessment of usability evaluation methods in an EHR. Int J Med Inf 2014;83(5):361-367.
24. Devine EB, Lee C, Overby CL, Abernethy N, McCune J, Smith JW, et al. Usability evaluation of pharmacogenomics clinical decision support aids and clinical knowledge resources in a computerized provider order entry system: a mixed methods approach. Int J Med Inf 2014;83(7):473-483.
25. Clarke MA, Belden JL, Kim MS. Determining differences in user performance between expert and novice primary care doctors when using an electronic health record (EHR). J Eval Clin Pract 2014;20(6):1153-1161.
26. Ozok AA, Wu H, Garrido M, Pronovost PJ, Gurses AP. Usability and perceived usefulness of personal health records for preventive health care: A case study focusing on patients' and primary care providers' perspectives. Appl Ergon 2014;45(3):613-628.
27. Corrao NJ, Robinson AG, Swiernik MA, Naeim A. Importance of testing for usability when selecting and implementing an electronic health or medical record system. J Oncol Pract 2010 May;6(3):120-124.
28. Mamykina L, Vawdrey DK, Stetson PD, Zheng K, Hripcsak G. Clinical documentation: composition or

synthesis? J Am Med Inform Assoc 2012 Nov- Dec;19(6):1025-1031.
29. Reeves S, Kuper A, Hodges BD. Qualitative research methodologies: ethnography. BMJ: British Medical Journal 2008;337.
30. Berg BL, Lune H, Lune H. Qualitative research methods for the social sciences. : Pearson Boston, MA; 2004.
31. 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.
32. Chen L, Guo U, Illipparambil LC, Netherton MD, Sheshadri B, Karu E, et al. Racing Against the Clock: Internal Medicine Residents' Time Spent On Electronic Health Records. Journal of graduate medical education 2016;8(1):39-44.
33. Altheide DL. Reflections: Ethnographic content analysis. Qualitative sociology 1987;10(1):65-77.
34. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004;24(2):105-112.
35. Zheng K, Padman R, Johnson MP, Diamond HS. An interface-driven analysis of user interactions with an electronic health records system. J Am Med Inform Assoc 2009 Mar-Apr;16(2):228-237.
36. Abran A, Khelifi A, Suryn W, Seffah A. Usability meanings and interpretations in ISO standards. Software Quality Journal 2003;11(4):325-338.
37. Nielsen J. Usability engineering. : Elsevier; 1994.
38. A toolkit for strategic usability: results from workshops, panels, and surveys. Proceedings of the SIGCHI conference on Human Factors in Computing Systems: ACM; 2000.
39. NASA-task load index (NASA-TLX); 20 years later. Proceedings of the human factors and ergonomics society annual meeting: Sage Publications; 2006.
40. Benedek J, Miner T. Measuring Desirability: New methods for evaluating desirability in a usability lab setting. Proceedings of Usability Professionals Association 2002;2003:8-12.