Comparing usability testing outcomes and functions of six electronic nursing record systems

Insook Cho [a,*], Eunman Kim [b], Woan Heui Choi [c], Nancy Staggers [d]

[a] Nursing Department, Inha University, Incheon, South Korea
[b] Department of Nursing, Sunmoon University, Chungnam, South Korea
[c] Department of Nursing, Seoul National University Hospital, Seoul, South Korea
[d] College of Nursing, University of Utah, Salt Lake City, USA

ARTICLE INFO

ABSTRACT

Objectives: This study examined the usability of six differing electronic nursing record (ENR) systems on the efficiency, proficiency and available functions for documenting nursing care and subsequently compared the results to nurses' perceived satisfaction from a previous study.

Methods: The six hospitals had different ENR systems, all with narrative nursing notes in use for more than three years. Stratified by type of nursing unit, 54 staff nurses were digitally recorded during on-site usability testing by employing validated patient care scenarios and think-aloud protocols. The time to complete specific tasks was also measured. Qualitative performance data were converted into scores on efficiency (relevancy), proficiency (accuracy), and a competency index using scoring schemes described by McGuire and Babbott. Six nurse managers and the researchers completed assessments of available ENR functions and examined computerized nursing process components including the linkages among them.

Results: For the usability test, participants' mean efficiency score was 94.2% (95% CI, 91.4–96.9%). The mean proficiency was 60.6% (95% CI, 54.3–66.8%), and the mean competency index was 59.5% (95% CI, 52.9–66.0). Efficiency scores were significantly different across ENRs as was the time to complete tasks, ranging from 226.3 to 457.2 s ($\chi^2 = 12.3$, $P = 0.031$; $\chi^2 = 11.2$, $P = 0.048$). No significant differences were seen for proficiency scores. The coverage of the various ENRs' nursing process ranged from 67% to 100%, but only two systems had complete integration of nursing components. Two systems with high efficiency and proficiency scores had much lower usability test scores and perceived user satisfaction along with more complex navigation patterns.

Conclusions: In terms of system usability and functions, different levels of sophistication of and interaction performance with ENR systems exist in practice. This suggests that ENRs may have variable impacts on clinical outcomes and care quality. Future studies are needed to explore ENR impact on nursing care quality, efficiency, and safety.

1. Introduction

Nurses worldwide are rapidly converting to electronic health records (EHRs) and electronic nursing records (ENRs) due to their projected benefits. According to a survey on trends in EHR adoption in non-federal acute care hospitals in the United States, 59% of hospitals adopted basic EHRs and more than 93% of these had certified EHR technology installed as of 2013 [1]. The rate of EHR adoption in the United Kingdom, Australia, and the Netherlands was 68%, 60%, and 33%, respectively; these rates were higher than in Canada (10%) and Germany (7%) [2]. A 2012 national survey used a definition of EHR functions similar to the U.S. and found that South Korea had an adoption rate of 45–95% for basic EHRs across type of hospital (primary, secondary, or tertiary) and 6.5–20% for comprehensive EHRs [3]. The survey also found that 70% of hospitals ($n = 202$) had electronic nursing record (ENR) systems with 10% having fully implemented ENRs with narrative nursing notes.

Increased adoption of EHRs has been accompanied by heightened recognition of usability issues. According to findings of physicians' satisfaction surveys from EHR users in the U.S., more than 30% of respondents would not recommend their EHR to oth-

* Corresponding author at: Nursing Department, Inha University, Inharo 100, Namgu, Incheon, South Korea.
  E-mail addresses: Insook.cho@inha.ac.kr, ICHO1@partners.org (I. Cho).

ers, and nearly 50% of respondents were not satisfied with many of the largest available EHR systems [4,5]. Several studies [6–8] also reported that poor EHR usability may result in user errors, untoward outcomes and unintended negative consequences. In some cases, user errors occurred without adverse events; in other cases, EHR-related system flaws led to adverse events or medical errors [7–10]. In response to the accumulating evidence, the Institute of Medicine and the American Medical Informatics Association recognized EHR usability as a critical component of safe and effective use of health IT. They made recommendations on enhancing patient safety and the quality of care, along with improving the usability of EHRs [7,11,12].

Despite the fact that usability is a major issue for health information technology (IT) adoption, little is known about the usability of ENRs. Several qualitative studies on nurses' perceptions of EHRs and ENRs showed mixed results [13–16]. A review of usability studies on health IT from 2003–2009 showed only 2 out of 346 studies evaluated health IT used by nurses [17]. These studies had many methodological issues, such as lack of a theoretical framework, lack of detail on qualitative approaches and a single-evaluation focus. In our previous national survey on the adoption of ENR systems at 289 hospitals in South Korea, we found wide variation in the nursing departments' satisfaction with ENRs with scores ranging from 16 to 95 out of a possible 100. The satisfaction scores had no relationship with type of hospital or staff composition [3]. As others have shown, this finding raised a concern that health IT, despite its great potential, may create new issues and have less than desirable outcomes. This concern motivated us to focus on ENR system-related issues and system usability for systems at different hospitals. We adopted the user–task–system interaction evaluation based on a stratified view of the health IT evaluation model suggested by Yen and Bakken [17].

In this study, we report usability testing data on six systems from six different hospitals as well as the relationships between usability test scores and previous satisfaction survey scores. The satisfaction survey findings were reported earlier [18]. Understanding usability issues and differences between systems can help us improve ENR systems supporting everyday clinical nursing practice.

2. Theoretical works

We found two published theoretical works in health informatics to be of potential relevance for our study: a stratified view of health IT usability evaluation (SV-HIT) [17] and the Task, User, Representation, and Function (TURF) framework [19]. SV-HIT authors reviewed and categorized 346 health IT usability studies into system development life-cycle stages to develop the SV-HIT. The model has three levels (Fig. 1). Level 1 targets system specifications to understand user–task interactions for a system development. The key questions for this level are "What are the needs/tasks?" and, "How can a system be used to support the needs/tasks?" Level 2 focuses on task performance to assess systems and human–computer interaction, a system–task or user–task–system type of evaluation. Sample key questions were, "Does the system work for the task?" or, "What is the user and system interaction performance, in terms of output quality, speed, accuracy, and completeness?" Level 3 incorporates environmental factors such as user variance, organizational support into a Level 2 evaluation to identify work processes and system impacts in real settings, a user–task–system–environment type of evaluation. The key questions include, "Is the user satisfied with the way the system helps perform a task in the real setting?" or "How does the system impact healthcare?"

TURF is based on work-centered research and an expansion of a previous framework called UFuRT developed earlier for the
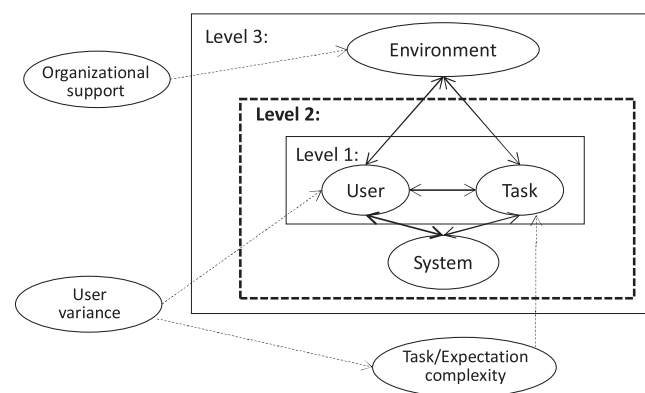

Fig. 1. Theoretical Base for EHR Usability Evaluation: a stratified view of health information technology usability evaluation (adopted from Yen and Bakken, [17]). The main focus of our study is indicated by a bold dotted line.

human–computer interaction field [20,21]. The framework was proposed to describe, explain, and predict usability differences in terms of a representation effect as well as to objectively define, evaluate, and measure usability. The framework focuses on the representation effect as the essence of usability. Although helpful, we think this framework is less relevant to the present study design compared to the SV-HIT so we used the SV-HIT to guide this study. Based on the SV-HIT model, we identified the present research as a Level 2 study examining user–task–system evaluation. We also analyzed system functions to get more information about each system. Specific purposes of the study were to (a) conduct a user–task–system analysis of ENRs on performance outcomes—relevancy and accuracy of data input and time consumed in data input, (b) conduct a functional analysis of ENRs and (c) examine the relationships between the performance outcomes and user survey findings.

3. Methods

A cross-sectional study was designed to evaluate ENR systems. The work was supported by the Korean Hospital Nurses Association (KHNA). The aim of the collaboration with KHNA was to examine ENR systems in practice and suggest improvements for these systems. This study was conducted with the approval of the institutional review board at each site.

3.1. Study sites and settings

Six general hospitals located in the Seoul metropolitan area and Gyeonggi province were invited to participate. These were all teaching hospitals with EHR and ENR systems (Table 1).

3.2. Nurse participants and usability testing tools

3.2.1. Participants

For the usability test, we recruited nine staff nurses from each hospital for a total of 54 clinical staff nurses. A stratified random sampling technique was used based on type of unit (medical, surgical and intensive care). Inclusion criteria were (a) nursing experience of more than 3 years on the unit, (b) experience of more than 3 years in the use of an ENR system and (c) willingness to participate. Initially, 108 nurses met the inclusion criteria. We randomly selected half or 54 nurses using a coin toss. The sample were all women and the average age was 27 years (SD, 2.67). The average years of acute care experience was 4.30 (SD = 2.75). The sample was nearly evenly divided into nurses with a bachelor's degree at 51.3% or a diploma degree at 48.7%. Nurses all worked

Table 1
Characteristics and ENR systems of participating hospitals.

Characteristics	Hospital					
	A	B	C	D	E	F
Number of beds	897	552	620	1792	1086	888
Length of EHR/ENR[a] use in years	5	3	5	5	4	5
Type of ENR	Locally developed	Vendor	Vendor	Locally developed	Vendor	Locally developed
Number of nurses involved in ENR management	None	1	None	1	1	1
Use of standardized nursing terminology	None	ICNP[b]	ICNP	ICNP	NANDA[c] only	ICNP, partial
Satisfaction score of the nursing department regarding ENR benefits; range: 0 – 100	50	75	85	96	75	75
Average number of patients cared for by a nurse on Medical and Surgical units	12–15	13	15–18	12–13	13	13–16
Average number of patients cared for by a nurse in Intensive care units	4	2–3	3–4	2–2.4	3	2–4

[a] ENR: electronic nursing record.
[b] ICNP: International Classification for nursing practice.
[c] NANDA: North American Nursing Diagnosis Association.

eight-hour shifts. We also invited one nurse manager from each hospital to participate in sessions on screen navigation patterns and the functional analysis; all were nurses who had been involved in the development of their ENRs.

3.2.2. Tools and scoring

We developed four patient scenarios using case studies from medical-surgical nursing textbooks. The first one was an inpatient scenario for a patient with a stroke and fever and a risk of pressure ulcers. We used this scenario for the system functional assessment. The other three scenarios, used for the usability test, included (a) a diabetic patient with pneumonia for nurses on medical units, (b) a patient having a kidney operation for nurses on surgical units and (c) a delirium patient with acute sepsis for nurses on surgical intensive care units. The scenarios were reviewed by two of the authors (I.C. and E.K.) for realism and assessed for comparable complexity as well as by a nurse manger with over 20 years practice in a medical-surgical nursing department in a hospital. The scenarios included patients' general information, a description of specific signs and symptoms, nursing care provided and communication with a physician.

Using a consensus-based approach, two reviewers developed criteria for desirable content for the associated nursing documentation: explicit assessments, actual nursing diagnoses, potential diagnoses, interventions provided and patients' responses. We used these criteria as gold standards for evaluating the nursing records entered by participants. Each scenario had a range of 13–22 sentences about diagnoses, assessments, and patient responses. If the user entered statements that exactly matched what should be documented according to the gold standards, the attempt was categorized as a "success," and a score of +1 was assigned. If the user entered phrases that did not match the gold standards, the attempt was categorized as "sub-optimal" and 0 was assigned. If the user could not enter expressions contained in the gold standards, the attempt was categorized as a "failure" and −1 was assigned.

We adopted the scoring mechanism described by McGuire and Babbott [22] for patient simulation exercises with nursing students. This consisted of an efficiency score, a proficiency score, and a composite index of overall competence. An efficiency score indicates the proportion of nurses' input that was deemed relevant to the scenario at hand. This is calculated by obtaining the total number of positive inputs and dividing that by the total number of all inputs made. The proficiency score represents the extent to which

the nurses' input was compatible with the defined gold standards. It is calculated by dividing the number of points earned by the maximum number of points possible. Finally, an overall competency index was calculated which combined the proficiency and efficiency scores into one index.

Functional analysis refers to the process of identifying a work domain's abstract structure, that is, the ontology of the work domain. Ontology is the basic structure of the work performed by the system together with its human users. The basic structure does not vary with respect to work procedures, application technology, or cognitive mechanisms [19]. We chose nursing process components as the nursing ontology for the functional assessment: assessment, diagnosis, outcomes identification, planning, implementation and evaluation. We also assessed the ENRs for available interactions such as drop-down menus and copy and paste features. Drop-down menus with predefined documentation options aid the organization of data entry for the patient information through the use of checklists and sequential structured entry of patient assessments, diagnoses, interventions provided and patients' responses. There were two kinds of predefined documentation options, which we referred to as nursing statements. One was nursing statements populated from and linked with standard terminologies available at hospital B, C and D. The other one was locally developed nursing statements without linkages to terminologies; most were in a free-text format at hospital A. Hospitals E and F had combinations of the two. All systems supported free-text input.

To compare the six ENR systems, we developed a method to determine functional coverage which refers to the number of nursing process components supported by a system. For example, if the narrative nursing notes supported the entry of each component of the nursing process, then the functional coverage was 100%. We also examined integration within the system, specifically with regard to (a) one-to-one linkages of nursing goals to specific interventions that are used to treat specific nursing diagnoses, and (b) the linkage of documented care planned and care delivered.

3.3. Procedures

After IRB approval, we collected data from August to November 2013 for the usability test and functional analysis.

3.3.1. Usability test procedure

The onsite usability tests were performed in a quiet conference room at each hospital. We used a portable video recorder and a microphone connected to a computer loaded with the site's installed EHR. We presented the appropriate written scenario to individual participants after they completed their shifts and then asked each to read the scenario and pretend to be a nurse taking care of the patient. After ensuring that the participants understood the scenario, we asked them to record nursing notes for the simulated patients using a test patient EHR account. To determine if nurses were able to accomplish the tasks, we requested they state loudly what they were looking at, thinking, doing, and feeling, as they performed the tasks. We videotaped the sessions, which lasted an average of eight minutes. The tapes were transcribed and annotated for efficiency and proficiency as described above. Two researchers (I.C. and E. K.) independently classified nurses' data entry. Discrepancies between the two reviewers were resolved by discussion. We also measured the time nurses took to document the nursing notes.

3.3.2. Comparison of usability test and user survey scores

The usability test results were compared with satisfaction survey scores reported in a previous study [18] that used Brooke's (1996) System Usability Scale (SUS) and a global satisfaction scale. In the previous survey, 1879 nurses participated from the same six hospitals in the present study.

3.3.3. Functional test procedures

We asked the nurse managers to explain how to use their ENR's narrative nursing notes and to demonstrate data entry based on the given scenario. We also gathered screen captures to illustrate each function. Two researchers (I.C and E. K.) independently checked each function and then calculated functional coverage by system. For any discrepancies or ambiguity in judging, several rounds of follow-up questions and requests for screen captures were used to resolve issues.

3.4. Data analyses

We analyzed the video recordings of the 54 nurses' interactions with the screens. We conducted screen navigation pattern analysis by drawing data input activities on the six hospital's mockup screens. Nurse managers' written explanations were also used in the analyses. Descriptive statistics were used to describe both the functional and usability testing outcomes. Scores on efficiency (relevancy) and proficiency (accuracy) were derived by comparing entries to the gold standards; the overall competency index were calculated as outlined above and then compared across sites using the Kruskal–Wallis test. An alpha level of 0.05 was set to test for significance.

4. Results

4.1. Usability test

Across the 54 nurses, the mean number of documentation entries was 12.4 sentences (range: 5–22 sentences). The overall efficiency (relevancy) of the nurse interactions across the six ENR systems was 94.2% (95% CI, 91.4–96.9%; range: 84.3–99.3%) as indicated in Table 2. The differences between the systems on efficiency were statistically significant ($\chi^2 = 12.29$, $P = 0.031$). The mean score for proficiency (accuracy) was 60.6% (95% CI, 54.3–66.8%; range: 42.8–68.6%). The overall competency indices ranged from 40.6% to 67.2% with a mean of 59.5% (95% CI, 52.9–66.0%). An average of 3 to 9.7 additional statements did not explicitly refer to material in the assigned scenario or were not relevant to the scenario. The nurses took a mean time of 317.2 s (95% CI, 273.7–360.8 s) to complete

Table 2
Scenario-based usability test scores of six narrative nursing notes.

Measurement	Total	Hospital						χ^2 (p)[a]
		A	B	C	D	E	F	
Efficiency, mean% (95% CI)	94.2 (91.4,96.9)	84.3 (73.3,95.4)	99.3 (98.2,100.0)	93.6 (88.6,98.5)	94.7 (90.7,98.8)	94.6 (89.2,100.0)	98.4 (95.4,100.0)	12.29(0.031)
Proficiency, mean% (95% CI)	60.6 (54.3,66.8)	42.8 (30.2,55.5)	64.4 (52.1,76.7)	57.2 (50.6,63.9)	68.6 (58.3,78.9)	63.5 (43.4,83.5)	66.6 (46.0,87.2)	10.27(0.068)
Competency index, mean% (95% CI)	59.5 (52.9,66.0)	40.6 (25.3,55.9)	64.1 (50.1,78.1)	55.7 (46.0,65.3)	67.2 (54.8,79.6)	62.8 (38.6,86.9)	66.1 (41.2,91.0)	10.93(0.053)
Additional entry, mean freq. (range)	5.04 (0–9)	2.8 (0–9)	5.2 (0–16)	3 (6–9)	4.6 (0–11)	9.7 (1–24)	5.0 (0–17)	NA
Time consumed, mean, s (95% CI)	317.2(273.7,360.8)	226.3(144.9,307.7)	252.5(192.9,312.2)	457.2(248.8,665.6)	361.4(306.8,416.0)	365.1(225.6,504.6)	305.8(224.9,386.6)	11.17(0.048)

[a] Kruskal–Wallis test.

the requested tasks; the times were significantly different across systems ranging from 226.3 to 457.2 s ($\chi^2 = 11.17$, $P = 0.048$).

Fig. 2 shows the navigation patterns that emerged from the video recording analysis. System A showed the simplest pattern as many entries were in a free-text format. Systems B, C, and D looked very similar with regard to screen designs and navigation patterns. The patterns of systems E and F were the most complicated.

4.2. A Comparison of usability test outcomes with previous usability survey findings

Fig. 3 displays the usability test results in the current study matched with the two scores of System Usability Scale (SUS) and perceived satisfaction. Systems A–D showed consistent relationships between the values; that is, systems with higher SUS and satisfaction scores also had higher scores on efficiency, proficiency as well as overall performance. Systems E and F showed much lower SUS and satisfaction scores despite higher scores on efficiency, proficiency and the overall performance index.

4.3. Functional analysis

All of the ENR systems had functions supporting nursing process components; however, the coverage ranged from 67% to 100%. One site did not have care planning functions (83%), and three did not have outcomes identification or care planning functions (67%). Two systems had all six functions (100%) (Table 3). As for system features, all systems supported "copy/paste," drop-down menus, and multiple inputs with check boxes.

Systems E and F supported sequential input of outcomes identification, care planning, and interventions mapped with a nursing diagnosis. System C included an option to designate links to specific interventions for documented diagnoses, but this is not mandatory. The rest of the systems did not require these explicit linkages. The integration of documented care planning and delivered care was available for Systems E and F only, supported by a pre-linked list. Of course, there was no way of establishing a link when a nurse used free-text input.

5. Discussion

To understand usability issues and differences among ENR systems, we conducted user–task–system interaction evaluations, completed a functional assessment and explored the relationship between usability test scores and scores on a previous usability survey. According to our findings, different levels of ENR system sophistication exist in practice. The assessed ENR systems have a high level of overall interaction performance but significant differences exist across systems. All have a low-to-moderate level of data accuracy with little variation across ENRs. Usability test results were not always consistent with previous user perceptions about usability, i.e., the degree to which systems are deemed useful, usable, and satisfactory. Navigation patterns indicated variations in complexity and in domain ontology for the nursing process. Outcomes identification and planning were the most frequently missed functions in systems. To our knowledge, this is a first quantitative study comparing ENR systems across hospitals.

5.1. Usability test

The average efficiency (relevancy) and proficiency (accuracy) scores were 94.2% and 60.6%, respectively. These findings mean that the systems were excellent for ease of data entry, but the quality of the entered data was relatively low. Five of the six ENRs showed equally high efficiency scores except for system A. System A was

associated with both the shortest efficiency and the lowest proficiency scores. This might be due to the characteristics of system A, which differed from the rest of systems because it featured free-text format and statements which were locally developed without nursing terminologies. This resulted in very simple user navigation patterns and the shortest input times.

Proficiency (accuracy) scores did not show any significant differences across systems despite the wide gap in scores. This was likely due to the large intra-system variation among nurses at each site. The variation was likely caused by other factors such as individual differences (user characteristics). Interestingly, the 95% confidence intervals for proficiency scores for systems E and F were much larger than the others, suggesting high variations in accuracy across nurses at these sites.

We can interpret the results in at least two ways. First, variation may be caused by participating nurses' characteristics such as age, clinical experience, and educational backgrounds. Effects of user variability are addressed in the SV-HIT EHR usability evaluation model. It is well known that users' cognitive abilities as well as specific EHR design features combine to create differences in users' system interactions [23]. Differences in expertise between more experienced nurses and less experienced nurses, consistent with Benner's model [24], also may have affected interactions with ENR systems.

For the system functional assessment, we assumed that nurses were all familiar with organizing knowledge of patient care into the nursing process. However, if this were true, nurses' interactions with ENR systems, in particular for proficiency, would have been much better than what was observed. The latter finding may imply that nurses have different abilities in organizing knowledge about practice into the structured methods of ENR documentation that uses the nursing process.

Second, the systems may have failed to reflect common needs of nurses as they record nursing notes. The five systems, except system A, supported the basic nursing process either in structured formats of pre-linked functions or by using checkboxes with lists of statements. For example, in some ENRs, selecting a nursing diagnosis was followed by outcome identification, planning, and interventions in a linear fashion. Therefore, selecting the right diagnosis was a key step for proficiency scores. In the future, systems may have a clinical decision support feature to recommend likely diagnoses; then, the gaps in proficiency scores might decrease. Systems E and F supported multiple methods of entering goals-plan-intervention statements simultaneously, resulting in wider variance in proficiency scores compared to the other systems. However, these features also caused more complex user interfaces and interaction navigation.

We observed additional information entered on routine care plans such as falls, pressure ulcers, and pain management regardless of the scenario or patients' identified risks. These safety indicators, which are included in the hospital accreditation program, are important. Thus, we noted it was difficult for nurses to adjust routine care plans into individualized ones. System E, for example, showed from 1 to 24 statements in routine care plans during one recorded session on the same scenario. Given the current ENR design, we observed a risk for patient care becoming less individualized.

5.2. Comparison of usability test results with user survey findings

A comparison of the current usability test results with previous user survey findings proved quite informative. Systems E and F had higher proficiency and efficiency scores but unexpected low scores on previous scores for usability and user satisfaction. The remaining four systems had a consistent relationship of scores across both usability performance and usability perceptions. This finding could
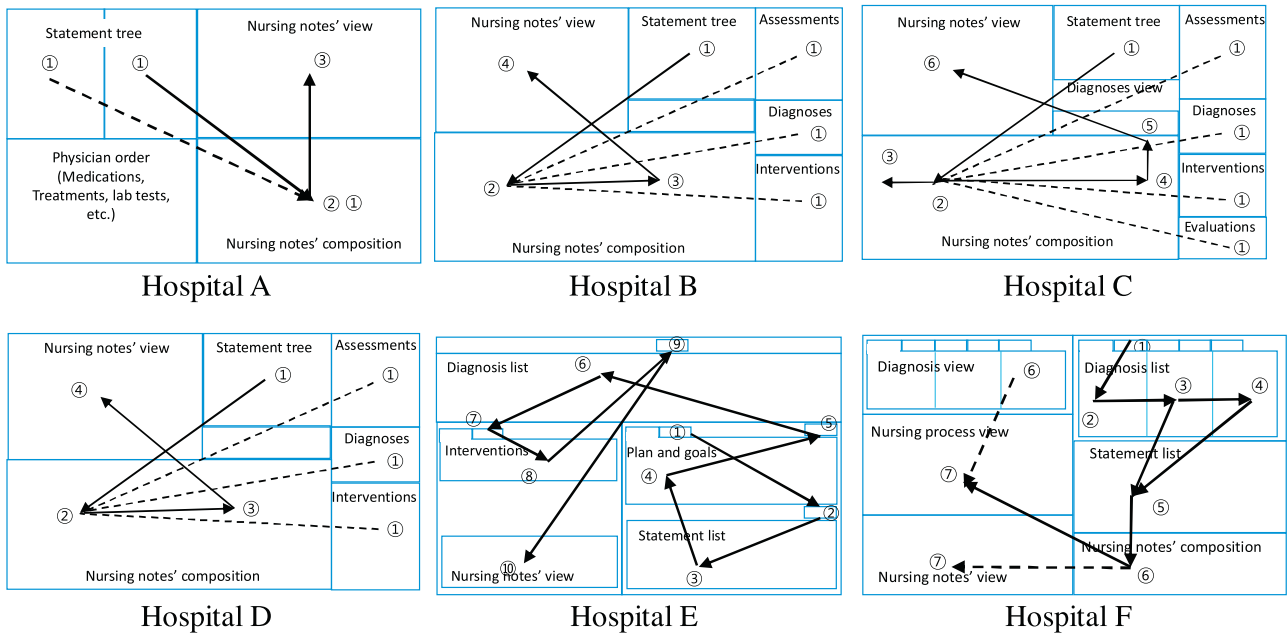
Fig. 2. User navigation patterns for the narrative nursing notes from the six systems. A solid line indicates the main pattern and a dotted line indicates an optional path. The circled digital numbers indicate common orders for data entry on each screen.
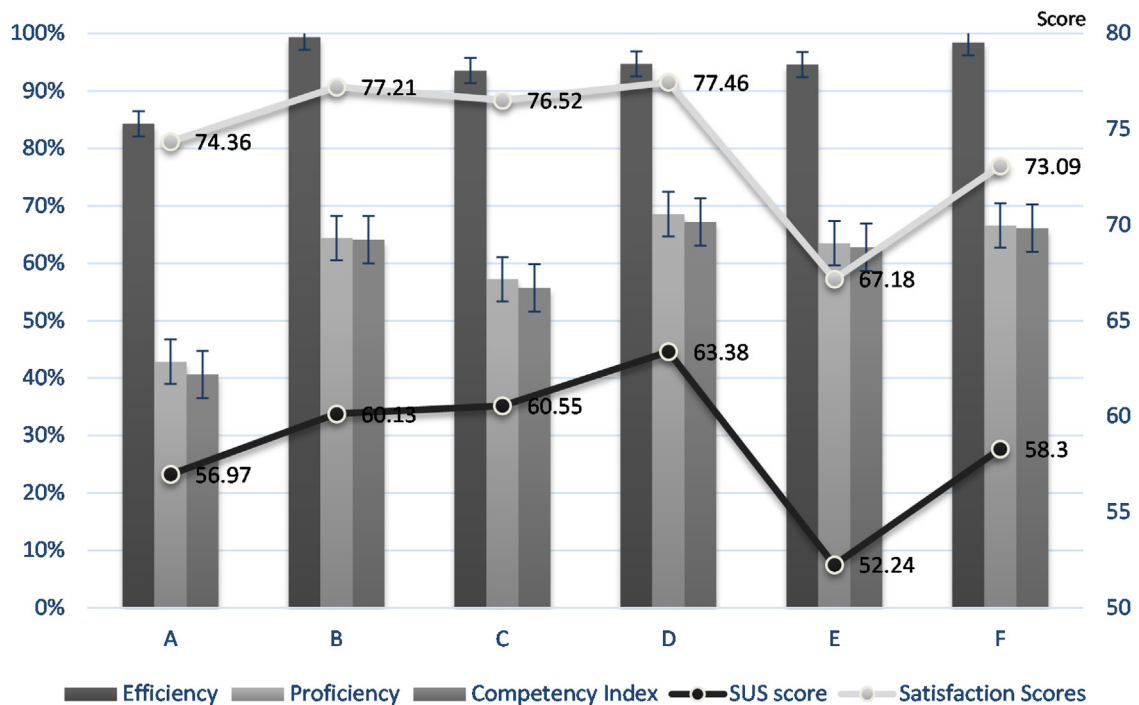

Fig. 3. A comparison of efficiency, proficiency, and the competency index, with the SUS scores and satisfaction scores of the six electronic nursing record systems.

mean that systems E and F were designed to support appropriate nursing documentation tasks, but the interaction design was too complex or the systems may not be a good fit with workflow in actual contexts. Nurses' "thoughtflow" may also differ from the designers' intent and user interface instantiation. In contrast, Systems B, C, and D had relatively high usability performance scores and user satisfaction, implying that the systems were more congruent with nurses' cognitive processes and designers' interface models. In the functional assessment, we found a hint of similar mismatches in systems E and F. These systems required outcomes identification, planning, and evaluation documentation in one lin-

ear direction. However, in practice, nurses employ bidirectional feedback loops from each component.

An interesting finding was that the three systems showing consistency between usability testing and survey scores also used standard nursing terminologies. This might be coincidence, or it might be because the use of standard terminology contributes to communication between nurses in a written format that reduces variations in expressions. A previous study [25] compared nurses' qualitative perceptions of the EHRs with and without nursing languages on documenting care. Nurses at a site without nursing language perceived more difficulty in using their EHRs than nurses

Table 3
Functional analysis of narrative nursing notes from six electronic nursing record systems.

Functions	Hospital[a]					
	A	B	C	D	E	F
Assessment	○	○	○	○	○	○
Diagnosis	○	○	○	○	○	○
Outcomes identification	○	⊗	⊗	⊗	○	○
Planning	⊗	⊗	⊗	⊗	○	○
Implementation	○	○	○	○	○	○
Evaluation	○	○	○	○	○	○
Coverage	83%	67%	67%	67%	100%	100%
Integration						
Linkage of a diagnosis-outcomes interventions	⊗	⊗	○	⊗	○	○
Linkage of a plan interventions	⊗	⊗	⊗	⊗	○	○

[a] The symbols ○ and ⊗ indicate the presence and absence of the relevant functions respectively.

at a site with nursing language. It was not clear whether the difference was only due to nursing language. Interestingly, our findings are consistent with this previous study. Further research is needed on how standard terminology might play a role in usability and user satisfaction.

Our comparisons allowed us to see clear differences between SV-HIT level 2 (user–task–system evaluation) and level 3 (user–task–system–environment evaluation) usability evaluations. User survey results showed that environmental factors, not included in our level 2 assessment, may play an important role in nurses' scores. Thus, a need exists to monitor how a system is used by end-users in actual settings and to determine specific contextual factors that might influence use. This issue also has been addressed elsewhere for EHR systems [26]. For example, the current increased rates of computerized provider order entry implementation for certified vendor products assumes improved patient safety will result. However, many studies suggested that system implementations are highly variable and can result in variable outcomes. Classen and Bates [26] recommended tools to evaluate systems after they are operational, not just before implementation. These tools would help ensure the meaningful benefits for ENR systems as well as assisting hospitals in ongoing self-assessment and improvement of their systems.

5.3. Functional analysis

Our findings showed that most systems have functions supporting the basic nursing process, but outcomes identification and care planning were most frequently missed. A previous extensive literature review focused on nursing diagnosis and care plans showed a lack of evidence about how nursing care plans affect patient outcomes and that care plans are seldom documented in clinical settings [27]. This was despite the fact that care plans are required by the Joint Commission on Accreditation of Healthcare Organization in the U.S. Recently, nursing care plans were crafted into evidence-based decision support systems within one integrated EHR where context- and patient-specific access to evidence were provided at the point of care [28]. The care plan functions available in systems E and F were traditional NANDA-based care plans and supported user selection of problems, goals, and activity plans in the form of a checklist. Nurses could easily enter care plan statements via multiple clicks, allowing automatic linkage to pre-entered diagnosis–goals-intervention plans. This approach looked like routine care plans and did not allow tailoring to individual patient conditions. Many nurses criticized ENRs for reducing their capabilities for individualized care because the predefined system choices were too broad to fit individual patients [13].

To help nurses control planning for the care process, the nursing process should be built on an informatics infrastructure that supports nurses' decision making and allows adjustments to care plans integrated into a EHR/ENR [28]. An example of how computers can take precedence over nurses' professional judgement is when an ENR gives prompts for patients at risk of developing pressure ulcers [29] despite patients not being at risk. Another instance is in inpatient units, where automatic prompts are used to remind nurses to note decompensation of a patient's vital signs [30]. A systematic review [31] of eight studies comparing computerized nursing care planning with paper-based systems indicated that computerization alone did not benefit nursing practice or patient care outcomes. Further research is needed on computerized care planning and linkages to informatics infrastructure for evidence-based practice to determine any derived benefits to nursing or patient care.

5.4. Limitations

Our study had several limitations. The onsite usability test employed working nurses without pre-selected individual characteristics. This may preclude generalization to other sites. Also, the time to complete tasks might be longer than the actual time in practice because of the think-aloud component of the test. Both of these might make it difficult to compare the results with other studies conducted in laboratory or real settings. However, nurses used simulated ENR systems and were experienced users of their systems. This proved to be very informative about how the ENRs were actually used in practice. Another limitation was that only six hospitals participated in the study; hence, the results might not be generalizable to all other hospitals such as community hospitals. The study did include six different ENR systems integrated into EHRs, and these were all installed for long periods of time. Therefore, these results are likely to be applicable to other similar hospitals in Korea as well as those of other countries with similar nursing practices, ENRs and processes. On the positive side, we stratified subjects according to nursing unit, and we used inclusion criteria based on clinical training and experience with ENR systems. The nine subjects from each site might not be representative of the whole institution, but even these small samples are sufficient to uncover usability issues. Authors indicate that five to eight users is a sufficient number to uncover most important usability issues [32,33].

Understanding usability issues and differences between systems can help ENR designers and implementers improve the design, implementation, and use of nursing-centric technology. Nurse managers, nursing informaticians, and coordinators should be able to compare their ENRs with others, to diagnose issues and to consider employing user-centered designs for the next generation of ENRs.

- Nurses worldwide are rapidly converting to electronic health records (EHR) due to their projected benefits, which has been accompanied by heightened recognition of usability issues.
- Poor EHR usability may result in user errors, untoward outcomes and unintended negative consequences.
- The usability of electronic nursing records (ENR) is not known well and several qualitative studies on nurses' perceptions of EHR and ENR showed mixed results.
- Understanding usability issues and differences between systems can help ENR designers and implementers improve the design, implementation, and use of the technology.

What does this study add to our knowledge?

- Different levels of ENR sophistication exist in practice in terms of system usability and functions.
- ENR systems have a high level of overall interaction performance and a low-to-moderate level of data accuracy with little variation across systems.
- Approached to nursing terminology, user-interface design and navigation strategy, and care planning functions were different across systems.
- Favorable results of usability test were not always consistent with nurses' perceptions of usability and user satisfaction.

6. Conclusions

In terms of system usability and functions, different levels of ENR sophistication were observed, and nurses' efficiency scores varied across sites. Favorable results for user–task–system evaluation were not always consistent with nurses' perceptions of usability and user satisfaction. These results suggest that an ENR's impact on clinical outcomes and care quality may be variable, too, especially if nurses' rely heavily on documentation for care continuity. For instance, differences in designs and depth of documentation could lead to information gaps or errors of omission and commission. To realize the promise of ENR on health care benefits, further study is needed to explore ENR impact on nursing care quality, efficiency, and safety.

Acknowledgements

References

[1] D. Charles, J. King, V. Patel, M.F. Furukawa, Adoption of electronic health record systems among US non-federal acute care hospitals: 2008–2012: Office of the National Coordinator for Health Information Technology (2013).
[2] C. Schoen, On the front line: primary care doctors' experiences in eleven countries. 2013: Findings from the Commonwealth Fund 2012 International Health Policy Survey of Primary Care Physicians and Health Affairs article, available from: https://xnet.kp.org/kpinternational/other/otherINTL.html.
[3] I. Cho, W.J. Choi, W.H. Choi, M.K. Kim, The adoption and use of electronic nursing records in Korean hospitals: findings of a nationwide survey, J. Korean Clin. Nurs. Res. 19 (3) (2013) 345–356.
[4] R.L. Edsall, K.G. Adler, The 2009 EHR user satisfaction survey: responses from 2,012 family physicians, Fam. Pract. Manag. 16 (6) (2009) 10–16.
[5] L.R. Kane, Electronic medical record survey results: medscape exclusive readers' choice, Medscape Bus. Med. (2009), Article 709856.
[6] J. Metzger, E. Welebob, D.W. Bates, S. Lipsitz, D.C. Classen, Mixed results in the safety performance of computerized physician order entry, Health Aff. 29 (4) (2010) 655–663.
[7] Institute of Medicine, Health IT and Patient Safety: Building Safer Systems for Better Care, The National Academies, 2011.
[8] J.M. Walker, P. Carayon, N. Leveson, R.A. Paulus, J. Tooker, H. Chin, et al., EHR safety: the way forward to safe and effective systems, J. Am. Med. Inform. Assoc. 15 (3) (2008) 272–277.
[9] A. Kushniruk, M. Beuscart-Zephir, A. Grzes, E. Borycki, L. Watbled, J. Kannry, Increasing the safety of healthcare information systems through improved procurement: toward a framework for selection of safe healthcare systems, Healthc. Q. (Tor., Ont.) 13 (2009) 53–58.
[10] B. Karsh, Beyond usability: designing effective technology implementation systems to promote patient safety, Qual. Saf. Health Care 13 (5) (2004) 388–394.
[11] D. Chou, Health IT and patient safety: building safer systems for better care, J. Am. Med. Assoc. 308 (21) (2012) 2282.
[12] B. Middleton, M. Bloomrosen, M.A. Dente, B. Hashmat, R. Koppel, J.M. Overhage, et al., Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA, J. Am. Med. Inform. Assoc. 20 (e1) (2013 Jun) e2–e8, amiajnl-2012-001458.
[13] J.E. Stevenson, G.C. Nilsson, G.I. Petersson, P.E. Johansson, Nurses' experience of using electronic patient records in everyday practice in acute/inpatient ward settings: a literature review, Health Inform. J. 16 (1) (2010) 63–72.
[14] R. Verwey, R.A.B. Claassen, M.J. Rutgers, L.P. de Witte, The implementation of an Electronic Nursing Record in a general hospital in the Netherlands: lessons to learn. Collaborative Patient Centered Ehealth, in: Proceedings of the HIT@ HealthCare 2008 Joint Event: 25th MIC Congress, 3rd International Congress Sixi, Special ISV-NVKVV Event, 8th Belgian EHealth Symposium, IOS Press, 2008.
[15] L. Poissant, J. Pereira, R. Tamblyn, Y. Kawasumi, The impact of electronic health records on time efficiency of physicians and nurses: a systematic review, J. Am. Med. Inform. Assoc. 12 (5) (2005) 505–516.
[16] J. Viitanen, H. Hypponen, T. Laaveri, J. Vanska, J. Reponen, I. Winblad, National questionnaire study on clinical ICT systems proofs: physicians suffer from poor usability, Int. J. Med. Inform. 80 (10) (2011 Oct) 708–725.
[17] P.-Y. Yen, S. Bakken, Review of health information technology usability study methodologies, J. Am. Med. Inform. Assoc. 19 (3) (2012 May–Jun) 413–422, amiajnl-2010-000020.
[18] I. Cho, W.H. Choi, M. Hyun, Y. Park, Y. Lee, E. Cho, et al., Identifying usability level and effecting factors of electronic nursing record systems: multi-Institutional time-motion approach, J. Korean Acad. Nurs. 45 (4) (2015 Aug) 523–532, http://dx.doi.org/10.4040/jkan.2015.45.4.523.
[19] J. Zhang, M.F. Walji, TURF: toward a unified framework of EHR usability, J. Biomed. Inform. 44 (6) (2011) 1056–1067.
[20] UFuRT: a work-centered framework and process for design and evaluation of information systemsJ. Zhang, K. Butler (Eds.), Proceedings of HCI International (2007).
[21] M. Nahm, J. Zhang, Operationalization of the UFuRT methodology for usability analysis in the clinical research data management domain, J. Biomed. Inform. 42 (2) (2009) 327–333.
[22] C.H. McGuire, D. Babbott, Simulation technique in the measurement of problem-solving skill, J. Educ. Meas. 4 (1) (1967) 1–10.
[23] B. Allen, Cognitive abilities and information system usability, Inform. Process. Manage. 30 (2) (1994) 177–191.
[24] P. Benner, From novice to expert, Am. J. Nurs. 82 (3) (1982) 402–407.
[25] J.M. Carrington, J.A. Effken, Strengths and limitations of the electronic health record for documenting clinical events, Comput. Inform. Nurs. 29 (6) (2011) 360–367.
[26] D.C. Classen, D.W. Bates, Finding the meaning in meaningful use, N. Engl. J. Med. 365 (9) (2011) 855–858.
[27] M. Müller-Staub, M.A. Lavin, I. Needham, T. Van Achterberg, Nursing diagnoses, interventions and outcomes-application and impact on nursing practice: systematic review, J. Adv. Nurs. 56 (5) (2006) 514–531.
[28] S. Bakken, L.M. Currie, N.-J. Lee, W.D. Roberts, S.A. Collins, J.J. Cimino, Integrating evidence into clinical information systems for nursing decision support, Int. J. Med. Inform. 77 (6) (2008) 413–420.
[29] I. Cho, I. Park, E. Kim, E. Lee, D.W. Bates, Using EHR data to predict hospital-acquired pressure ulcers: a prospective study of a Bayesian Network model, Int. J. Med. Inform. 82 (11) (2013) 1059–1067.
[30] H. Brown, J. Terrence, P. Vasquez, D.W. Bates, E. Zimlichman, Continuous monitoring in an inpatient medical-surgical unit: a controlled clinical trial, Am. J. Med. 127 (3) (2014) 226–232.
[31] C. Urquhart, R. Currell, M.J. Grant, N.R. Hardiker, Nursing record systems: effects on nursing practice and healthcare outcomes, Cochrane Libr. (2009).
[32] A mathematical model of the finding of usability problemsJ. Nielsen, T.K. Landauer (Eds.), Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems (1993), ACM.
[33] J.R. Lewis, Sample sizes for usability studies: additional considerations, Hum. Factors: J. Hum. Factors Ergon. Soc. 36 (2) (1994) 368–378.