Published in final edited form as:
Int J Med Inform. 2012 December ; 81(12): 842–851. doi:10.1016/j.ijmedinf.2012.04.004.

Evaluation of an integrated graphical display to promote acute change detection in ICU patients

Shilo Anders, PhD[1,2,3], **Robert Albert, MA**[8], **Anne Miller, PhD**[1,2,4], **Matthew B. Weinger, MD**[1,2,3,7], **Alexa K. Doig, PhD, RN**[6], **Michael Behrens, MD**[5], and **Jim Agutter, MArch**[5,8]

[1]Center for Research and Innovation in Systems Safety, Vanderbilt University School of Medicine, Nashville, TN

[2]Department of Anesthesiology, Vanderbilt University School of Medicine, Nashville, TN

[3]Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN

[4]School of Nursing, Vanderbilt University School of Medicine, Nashville, TN

[5]Department of Anesthesiology, University of Utah, Salt Lake City, UT

[6]College of Nursing, University of Utah, Salt Lake City, UT

[7]Geriatric Research Education and Clinical Center, Veterans Affairs Tennessee Valley Healthcare System, Nashville, TN

[8]Medvis, LLC, Salt Lake City, UT

Abstract

Objective—The purpose of this study was to evaluate ICU nurses' ability to detect patient change using an integrated graphical information display (IGID) versus a conventional tabular ICU patient information display (i.e. electronic chart).

Design—Using participants from two different sites, we conducted a repeated measures simulator-based experiment to assess ICU nurses' ability to detect abnormal patient variables using a novel IGID versus a conventional tabular information display. Patient scenarios and display presentations were fully counterbalanced.

Corresponding Author: Shilo Anders, PhD, Center for Research and Innovation is Systems Safety, 1221 21st Avenue South, Medical Arts Building, Suite 732, Nashville, TN 37212, (615) 936-3614, Fax (615) 936-7373, Shilo.anders@vanderbilt.edu.

Authors' contributions

Shilo Anders contributed to the study design, acquisition of data, analysis and interpretation, drafts and revisions of the article, and final approval for submission. Robert Albert contributed to the study design, acquisition of data, analysis and interpretation, drafts and revisions of the article, and final approval for submission. Anne Miller contributed to the study design, acquisition of data, analysis and interpretation, drafts and revisions of the article, and final approval for submission. Matthew B. Weinger contributed to the study design, drafts and revisions of the article, and final approval for submission. Alexa K. Doig contributed to the study design, acquisition of data, analysis, and final approval for submission. Michael Behrens contributed to the study development and final approval for submission. Jim Agutter contributed to the study design, acquisition of data, analysis and interpretation, drafts and revisions of the article, and final approval for submission.

Conflict of interest statement

Jim Agutter has an equity interest in MedVis LLC which has intellectual property rights to the graphical display. None of the other authors have any interests in MedVis or the technology.

Measurements—We measured percent correct detection of abnormal patient variables, nurses' perceived workload (NASA-TLX), and display usability ratings.

Results—32 ICU nurses (87% female, median age of 29 years, and median ICU experience of 2.5 years) using the IGID detected more abnormal variables compared to the tabular display [$F(1,119)=13.0$, $p < 0.05$]. There was a significant main effect of site [$F(1, 119)=14.2$], with development site participants doing better. There were no significant differences in nurses' perceived workload. The IGID display was rated as more usable than the conventional display, [$F(1, 60)=31.7$].

Conclusion—Overall, nurses reported more important physiological information with the novel IGID than tabular display. Moreover, the finding of site differences may reflect local influences in work practice and involvement in iterative display design methodology. Information displays developed using user-centered design should accommodate the full diversity of the intended user population across use sites.

Keywords

medical informatics; human engineering; man-machine systems; intensive care

II. INTRODUCTION

The nature of health care delivery is changing. In the future, prevention-oriented care will keep patients out of hospitals and many patients currently in hospitals will instead be treated in short-stay community settings (1). Hospitalized patients are likely to be more complex and more susceptible to complications. In this context, effective, high-quality care will depend on rapid and accurate detection of clinical cues of possible complications or adverse events. Patient information that is integrated and presented in an electronic system, so as to facilitate practitioners' ability to identify potentially adverse shifts and patterns may decrease the risk of complications. The purpose of this study was to test the effects of a trending system using an integrated graphical information display (IGID) that logically organized disparate patient information along a time-line in a single contextual graphic display.

Information display format can influence practitioners' ability to identify potentially adverse changes in physiological data (2). Ash et al found that medical displays are often incompatible with practitioners' workflow and unnecessarily fragment patient information (3). Information is often spread across multiple tabs and locations that require piecemeal information search and acquisition. This may confound practitioners' ability to detect evolving changes, make it more difficult to attain a holistic view of a patient's state, lead to care inefficiencies, and frustrate clinicians.

Miller, Scheinkestel and Steele found that displays that integrated patient information in physiologically meaningful ways better supported ICU nurses' ability to detect changed parameters (4). Their prototype was based on the following three design goals: 1) grouping related parameters together and separating unrelated parameters; 2) showing cause-and-effect relationships (e.g. showing percentage of inspired oxygen next to oxygen saturation) and 3) presenting patient data in the context of treatment goals (e.g. desired blood pressure ranges).

Concurrently, others on the research team developed integrated graphic displays for anesthesiologists (2, 5–8). The design goals were decreased response time, improved situation awareness, and better adverse event detection in anesthetized patients. In a scenario involving heavy fluid loss, anesthesiologists responded faster when using the integrated

graphic display than a traditional waveform display (5). In contrast, there were no differences between displays in a transfusion reaction scenario. Thus, there is ample evidence that properly designed, contextually relevant integrated graphical displays could be beneficial in a range of acute care environment (9, 10).

Intensive care unit (ICU) patients are usually dependent upon life-sustaining supportive therapy. In this environment, nurses' ability to detect and assess changing patient states is a first-line of defense against deteriorating condition or adverse events. In addition to their many other care responsibilities (medication administration, etc), nurses monitor patients' physiologic states and communicate any concerns to physicians and other healthcare providers. ICU patients may deteriorate rapidly (even suddenly) or very slowly over time (days). Both situations are critical to detect and accurately diagnose so that timely treatment can be administered.

In the present study, we compared interactive prototypes of an IGID with a more conventional tabular display to examine ICU nurses' ability to communicate identified abnormal patient parameters. We additionally wanted to ensure that the effects were equivalent across study sites.

The study was designed to test the following hypotheses:

Hypothesis 1: Nurses using the IGID will detect and communicate to a physician (confederate) more abnormal patient physiological states than when using the tabular (control) display.

Hypothesis 2: Nurses using the IGID will report lower psychological workload than when using the tabular display.

Hypothesis 3: Nurses using the IGID will be more satisfied than when using the tabular display.

III. METHOD

A. Context

The study was undertaken at two different academic medical centers. The site profiles are shown in Table 1.

B. Participants

Following IRB approval at both hospital sites, 32 volunteer ICU nurses (16 at each site) responded to email and flyer-based calls for participation. All participants were screened and excluded from participation if they had previously participated in any of the graphical display design and feedback sessions. Of the 16 nurses at Site A, 12 were female and all had up to 3 years of ICU work experience. Of the 16 nurses at Site B, 14 were female and 11 had up to 3 years ICU work experience. This group was initially selected so that expertise bias associated with diagnosing critical ICU events would be minimized. Due to a paucity of ICU nurses with 3 years or less experience at Site B, 5 additional nurses with between 3.5 and 18 years of experience were accepted as study participants, as a result of this our analysis controlled for experience.

C. Materials

Materials used to test our hypothesis included two patient scenarios with a set of Delphi scores, two patient data displays, the NASA-TLX and a pre-/post-test usability survey. Each of these materials is described below.

1. Design of ICU Displays—The graphic display design was based on the extensive display design experience of the team, literature evidence, and clinician input. The final IGID incorporated both trended patient information as well as patient state configural information displays (5). The display design addressed three themes identified during the research phase: display all relevant patient information on one screen; Highlight patient specific trending information so that a care provider could assess overall patient status and trajectory; and organize the information based on variables associated with a system rather than a measurement or testing device. We employed user-centered design (11, 12) with iterative user feedback cycles as described previously (5–8). Clinician input and the majority of design tradeoff decisions occurred at Site A where the development team was located.

As can be seen in Figure 2, all patient information, including medication administration, events and vital signs, was grouped functionally in panels that could be expanded or contracted at the users' discretion. This approach is novel in that it allows the user to coordinate relevant information across time. Traditional displays often display this information in a number of windows, thus relying on the user to cognitively integrate this information. Any variable that was beyond the target or goal range was highlighted in red based on accepted 'normal ranges.' However, clinicians could change these thresholds depending on the patient's disease trajectory and clinical objectives (4). When contracted, the names and current values of the variables were visible and out-of-range variables highlighted on the group bar. At the top of the display, a scalable timeline allowed clinicians to move their focus of attention and 'zoom in' on specific data points or 'zoom out' to see extended timeframes.

At the bottom of the display, two configural graphic elements portrayed an integrated collection of cardiovascular variables and arterial blood gas (ABG) variables. The cardiovascular graphic was previously designed and tested (5, 6). The ABG graphic was developed using the same UCD techniques specifically for this study.

The study sites used different commercial electronic ICU charting systems that reflected former tabular-numeric paper-chart designs. Thus, to have a standard comparative display, we developed a generic emulation of these traditional tabular-numeric systems for our control condition. For this interface, patient data were displayed as a table of numerals. The display had four screens that participant's viewed to obtain patient information such as vital signs, fluid inputs and outputs, procedures, and drugs. As shown in Figure 2, the vitals tab contained information about various lab test results, vital signs, and ventilator setting variables. The inputs (ins) and outputs (outs) tab showed the patient's fluid balance over time, the procedures tab indicated all patient procedures, and the medications tab showed the type, dose, time, and who administered any medications. Both displays always contained the same clinical data for the scenario being presented, which included all of the relevant information to achieve a perfect performance score.

2. Patient scenarios and Delphi scores—Domain experts (MB, AM, AD) developed three patient scenarios: one for training and two for testing based upon relevance to the ICU environment. The training scenario involved two parts each with a different pathology: failed weaning trial resulting in respiratory acidosis followed by renal failure. The test scenarios also had two parts with different pathologies. Each part contained three segments of the developing pathology over time for a total of six segments in each test scenario. An initial written patient history complemented the ICU patient data shown in the display. These data changed over time to match the pathology's physiologic changes depending on which segment was being presented and included physician orders. The patient in the first test scenario developed sepsis over the first three segments and abdominal compartment

syndrome over the last three. The second patient's pathology included septic shock from bowel perforation, followed by a transfusion reaction.

Scoring metrics for each scenario segment were developed independently using the modified Delphi process (13, 14) by a team of six Site A domain experts (i.e., 2 ICU physicians, and 4 critical care nurses) and one Site B ICU physician. The patient data in each scenario segment were reviewed separately by the experts who then rated each item on a 5-point Likert scale, (15) where 0 represented 'this parameter is not a critical indicator' and 5 represented 'this parameter is a critical indicator that should be communicated by a nurse to a physician'. The experts were also encouraged to add new items to the list. Prior to each iteration experts were provided a summary of the mean and median ratings from the previous iteration. The experts' scores converged after three survey iterations and this became the consensus checklist of information that should be communicated by a nurse to a physician for each scenario segment.

For example, a scenario where the patient was presenting with sepsis, the domain experts rated urine output low as 5, which was included in the scoring metric. Conversely, in this instance, experts rated blood pressure as 0, thus it was excluded from the scoring metric. The scoring metric for each part of the scenario contained between 13 and 32 pieces of patient information deemed critical to communicate. The performance score was the proportion of items reported by the nurse to the physician.

3. NASA-Task Load Index—The National Aeronautics and Space Administration (NASA) Task Load Index (TLX) is the most widely used multidimensional scale to assess subjective workload across a variety of domains and activities (16). More recently, it has been used to assess workload associated with user interface designs (5, 17). The NASA-TLX consists of 6 subscales weighted on their individual significance for each participant: mental, physical, and temporal demands, frustration, performance, and effort. We administered the NASA-TLX after each scenario. The resulting data were analyzed in the standard manner as described by Hart and Staveland (16).

4. Pre- & Post-Test Survey—The 21 item pre-test survey consisted of a 7-point Likert-type scale rating for each statement and three open-ended questions about the current electronic charting systems in the ICU where a participant worked. The survey consisted of statements that addressed ease of use, usefulness, satisfaction, and support of situation understanding (Table 2). Usefulness statements assessed display navigation and functionality, while ease of use statements explored user's perceived ease of understanding and interpretation of the information. Satisfaction statements assessed nurses' likely use in actual patient care and situation understanding statements assessed the display's ability to present information that is easily integrated and understood. The post-test survey was identical to the pre-test survey with the addition of four configural graphics questions.

D. Study design/Procedure

Each participant was tested individually in a single session that lasted from one hour and forty-five minutes to three hours. Participants first completed informed consent, a basic demographic questionnaire, the pre-test survey, and the NASA-TLX weighting component. The participant used a desktop computer that presented each interface and was provided with a telephone. All participants were audio recorded during the session.

Participants were trained on the components of each display culminating in a knowledge demonstration test and practice scenario in which each display was used. For the practice scenario, participants were instructed to review the patient's displayed data and to use the telephone to initiate communication with the patient's physician as soon as a reason to call

became apparent. The practice scenario included use of both displays. The presentation of the displays and scenarios were counterbalanced across all participants to minimize learning and order effects.

After training, the participants were presented with a brief patient description, including chief complaint, prior history, pertinent findings, and medications. Participant questions were answered before proceeding to the first test scenario. The participants used the display to explore the patient's data, and initiated a phone call to the confederate clinician whenever, in their clinical assessment, the situation required. If no call was initiated by the nurse within two minutes, the physician would call requesting a patient update. A scenario-specific checklist was used to score the content of the nurse's verbal report. Once the participant had given the report, the phone call ended and a word search distracter task was completed for two minutes prior to beginning the next segment of the scenario. Each scenario had two parts as described above, which were included in the analyses.

After the first patient scenario was completed, the participant completed a NASA-TLX and the post-test survey. The second scenario used the same procedure with the other display.

E. Data analysis

To minimize the Delphi performance scoring differences, raters from both sites met to analyze all of the participants' conversations using the previously developed scenario-specific scoring checklists. Reviewers were blinded to the extent possible as to display condition. Differences in scoring between reviewers were resolved by consensus.

The experimental design of the performance scores for this study was a within-subject comparison of display type (IGID and tabular) and between-subject comparison of location site (Site A and Site B). Each of the two test scenarios was divided into two tasks, because each part presented a different pathology, set of reportable patient issues and performance data points. Thus, for every participant there were 2 performance data points for each display. A two-factor analysis of co-variance (ANCOVA) test was used to analyze the performance scores, which controlled for participant experience and test scenario. The dependent variable was the percent correct on the Delphi score.

We conducted a two-factor analysis of variance (ANOVA) for display type and site differences on the NASA-TLX rating score. Traditionally, responses to a single Likert-type item are treated as ordinal data in which case a non-parametric test is appropriate; however, summed responses are often treated as interval data to measure variablility. Thus, the total and grouped usability questions were analyzed using two-factor ANOVA with display condition and site as the independent variables and the dependent variable was the sum of several 7-point Likert item responses. All data were analyzed using SPSS 19. Using a Bonferroni correction for family-wise error (20), was defined as a P value of less than 0.0167 (two-tailed test).

III. RESULTS

Demographics

A total of 32 ICU nurses from two sites participated in this study. 87% of the participants were female with a median age of 29 years (range 21–52). The median ICU nursing experience was 2.5 years (mean 3.2 years). The majority of the ICU nurses worked in a medical ICU (68.8%), but included surgical (6.2%), cardiovascular (9.4%), burn (9.4%), and other (6.2%). 14 participants at Site A routinely used ICU EHR software from CERNER, while 17 at Site B used McKesson Horizon Expert Documentation system (HED). One participant used neither documentation system.

Delphi

There was significantly better performance (Figure 3) using the graphic display (μ=52.2%) compared to the tabular display (μ=45.6%) [F (1,119)=13.0, p < 0.0167, used for all analyses]. Further, there was a significant main effect of site [F (1, 119)=14.2] in which Site A (μ=53.6%) performed better than Site B (μ= 44.1%).

Moreover, there was a significant site-by-display type interaction [F (1, 119)=6.4]. While the two sites' participants did not differ in their performance when using the tabular display [F (1, 61)=1.4], Site A performance was much better in the graphic condition [Site A 59.2% ± 10.1 versus Site B 45.1%±12.1, F (1, 61)=16.8]. A condition-by-scenario ANCOVA controlling for experience confirmed the results of the primary analysis with a significant difference between display conditions at Site A [F (1, 59)=20.7] but not Site B.

Workload

Workload was measured using the composite score of the weighted NASA-TLX where a lower score corresponded to lower perceived workload. The workload score reported by participants for the display they currently used in their ICU was μ = 59.0±13.5. After the scenario using the IGID, workload was μ=52.8±14.4, while workload reported after using the tabular display was μ=56.5±12.7. There were no statistically significant differences of the site, display condition or an interaction effect.

Usability Survey

Overall, the IGID was rated significantly higher for its usability (μ=6.0±0.5) than the display they currently used in practice [μ= 5.0±0.9, F (1, 60)=31.7, p<0.0167]. There were no differences in usability ratings by site nor a site-by-display interaction. Each sub-group of the usability survey, rather than individual item was further analyzed using a paired t-test. As can be seen in Figure 4 the usability questionnaire sub-groups were all rated significantly higher in the IGID display than the current display [Usefulness, t(62)=4.9, p<0.0167; Ease of use, t(62)=3.3, p<0.0167; Satisfaction, t(62)=2.8, p<0.0167; Situation understanding, t(62)=7.6, p<0.0167].

IV. DISCUSSION

In this study, we employed an iterative UCD technique to develop an integrated graphical trending display intended to better support ICU nurses' ability to synthesize information about patients' evolving clinical conditions. The overall results supported our hypothesis that ICU nurse participant's performance was better using the IGID. However, there was a site difference with the nurses performing much better at Site A. All study participants subjectively preferred the graphic display; however, their perceived workload was similar using either display. These findings have implications for the generalizability of user interface designs optimized for use at one clinical site as well as the conduct and interpretation of clinical usability tests with both functional (e.g., correct diagnosis) and subjective metrics (e.g., preference surveys).

These results are similar to previous clinical studies evaluating integrated displays (5–8, 21). A recent systematic review of other novel graphic displays found that in most studies the detection of adverse events improved as did clinical decision or diagnosis accuracy (22). Display designs of pulmonary, hemodynamic, and integrated anesthesia variables have shown significantly improved accuracy and/or speed of performance. While prior studies examined performance with brief clinical "snapshots", the present study allowed ICU nurses to evaluate an evolving patient event that occurred over several hours and measured

performance not by event recognition, but in what physiological data were deemed meaningful to report to confederate physicians.

Interestingly, nurses from Site A performed significantly better using the IGID than at Site B. This may be due to a number of factors (see Table 3), indicating that further study is needed. One contributing factor may be that Site A provided more feedback during the UCD resulting in a system that reflected local preferences. This may have had a disproportionate effect that favored performance at Site A more than at Site B. Additionally, this finding seems similar to the findings of a literature review on clinical decision support systems that suggests those systems where an author was involved in the development resulted in improved performance when compared with those where the development had been independent (20).

The IGID was designed for flexibility and configurability which allowed for a multitude of information arrangements to fit with different clinicians and site needs. The default functional groupings of the physiological variables were based initially on Miller et al (4) but were modified in some cases based on Site A clinical user preferences. For example, mean arterial pressure (MAP) was positioned between systolic and diastolic blood pressures, but local preference placed it beneath the diastolic. These rearrangements based on local preferences may have had an effect on performance at the other site. Although the ICU nurses level of experience was controlled as a covariate in our analysis, environmental characteristics like training and use of standardized handover frameworks (e.g. SBAR), and attributes of currently used electronic health records (which were different in each site) may have contributed to site differences. The table below summarizes the potential reasons for the site differences.

The results may under-estimate overall performance. Our intention was to standardize the experimental context as much as possible across the sites. Rather than compare performance using the IGID against the different local information systems, we constructed a comparative system that was a hybrid of both local displays that was an unfamiliar to all participants thus eliminated any training effects of display use. While this improved the fairness of the comparison, participants had to learn both displays. As a result, the highest mean performance was less than 60% correct. This low performance may have been a result of the experimental study context, while the use of clinical scenarios contributed to a more realistic test setting, the actual context of being in the ICU, having a real patient, other ICU tasks, and additional technology use may affect. Another potential reason for low performance scores may be reflective of the current communication between nurses and physicians in the ICU. Previous studies have pointed to poor communication resulting in less than optimal patient care. (3, 21). Finally, prior team training in these environments may have led the nurse to communicate the least amount of required information to incite the physician to action (e.g. go to the ICU). Future studies should include more contextually based performance tests, with use in the actual ICU or with high-fidelity simulation.

There are several limitations to the current study. The evaluation did not study IGID use in nurses' actual work in a realistic environment. In a clinical environment providers would have access to far more information from the actual patient, colleagues and other sources of documentation. While every attempt was made to make the IGID a central repository of all patient related information, further study is needed to see how this approach may be clinically significant within the clinical environment and integrate with other ICU technologies (e.g. bedside monitors), which was not the purpose of this study.

This study presents implications related to health information technology design and evaluation. A multi-site iterative development process with substantial contributions from

diverse sites may yield more generalizable results. The general interface components (e.g., functional groupings) of a display may be universal, but other components (e.g., labeling and listing conventions, the grouping of variables that overlap physiological systems) may be site or type of site specific. Therefore, system developers may want to allow for flexibility and configurability based upon site needs and clinician expertise and job function.

While user perceptions and other subjective data has clear value, our results suggest the need for caution in their interpretation. Participants consistently preferred the graphical display in the absence of feedback about their performance in the simulated studies. We do not know how performance feedback would have affected survey responses. In real-world environments, users receive direct and indirect performance feedback, when using technology, and this can effect their perceptions, acceptance, and adoption. More research on this topic is warranted.

Sanderson has suggested that there is insufficient data about how integrated displays affects clinicians' mental workload (10). In a recent meta-analysis of evaluations of physiological monitoring displays, less than half of the studies using the NASA-TLX revealed significant differences between displays (22). Alternative methods for measuring cognitive load and stress may need to be applied in information perception and processing studies.

V. CONCLUSION

This study suggests that an ICU patient information display using a graphic user interface improved nurse performance when compared to a tabular display. However, the finding of a site performance difference raises questions that necessitate further study with guided hypotheses. These include the effects of local influences on design evaluation, information system usage and generalizability. The data suggest the need for additional healthcare data display testing with diverse users and in diverse locations. Formal usability testing in simulated clinical settings can be expected to provide key insights into the design of electronic health records and other HIT. There are several limitations to the current study. These include the use of a simulated environment, representative cases and a composite control system. The results present nursing performance data using the interface in a non-ICU setting, and thus cannot necessarily be extrapolated to other practitioners, as well as performance in actual clinical settings where HIT use must be integrated into patient care processes. Additionally, the tabular display might not be analogous to the nurse's current tabular display and as such not indicative of actual performance. These data suggest that one size may not fit all, and that medical information system developers should build in configurability and flexibility to allow for local preferences.

Acknowledgments

Thanks to the nurses and physicians who provided feedback during design of the graphical interface and who participated in the development of the performance measure for the study. This research was supported by a National Heart, Lung and Blood Institute SBIR Grant (5R43HL079783-02).

VI. REFERENCES

1. Christensen, C.; Grossman, J.; Hwang, J. The Innovator's Prescription: A Disruptive Solution for Health Care. New York: McGraw-Hill; 2009.
2. Effken J, Loeb R, Kang Y, Lin Z. Clinical information displays to improve ICU outcomes. Int J Med Inform. 2008; 77:765–77. [PubMed: 18639487]
3. Ash J, Berg M, Coiera E. Some unintended consequences of information technology in health care: The nature of patient care information system-related errors. J Am Med Inform Assoc. 2004; 11(2): 104–12. [PubMed: 14633936]

4. Miller A, Scheinkestel C, Steele C. The effects of information presentation on physicians' and nurses' decision-making in ICUs. Applied Ergonomics. 2009; 40(4):753–61. [PubMed: 18834970]
5. Agutter J, Drews F, Syroid N, Westenskow D, Albert R, Strayer D, et al. Evaluation of a graphic cardiovascular display in a high fidelity simulator. Analg Anesth. 2003; 97:1403–13.
6. Albert R, Agutter J, Syroid N, Johnson K, Loeb R, Westenskow D. Simulation based evaluation of a graphic cardiovascular display. Anesth Analog. 2007; 105:1303–11.
7. Syroid ND, Agutter J, Drews FA, Westenskow DR, Albert RW, Bermudez JC, et al. Development and evaluation of a graphical anesthesia drug display. Anesthesiology. 2002; 96(3):565–75. [PubMed: 11873029]
8. Wachter SB, Agutter J, Syroid N, Drews F, Weinger MB, Westenskow D. The employment of an iterative design process to develop a pulmonary graphical display. J Am Med Inform Assoc. 2003; 10:363–72. [PubMed: 12668693]
9. Blike G, Surgenor S, Whalen K. A graphical object display improves anesthesiologists performance on a simulated diagnostic task. J Clin Mon Com. 1999; 15:37–44.
10. Sanderson P, Watson M, Russell W. Advanced patient monitoring displays: Tools for continuous informing. Anesth Analog. 2005; 101:161–8.
11. Hackos, J.; Redish, J. User and Task Analysis for Interface Design. New York: Wiley; 1998.
12. Rubin, J. Handbook of Usability Testing. New York, NY: John Wiley & Sons; 1994.
13. Clayton M. Delphi: A technique to harness expert opinion for critical decision-making tasks in education. Edu Psych. 1997; 17(4):373–86.
14. Scavone B, Sproviero M, McCarthy R, Wong C, Sullivan J, Siddall V, et al. Development of an objective scoring system for measurement of resident performance on the human patient simulator. Anesth. 2006; 105:260–6.
15. Clason D, Dormody T. Analyzing data measured by individual Likert-type items. J Agricultural Edu. 1994; 35(4):31–5.
16. Hart, SG.; Staveland, LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, PA.; Meshkati, N., editors. Human Mental Workload. North-Holland: Elsevier Science Publishers; 1988. p. 139-83.
17. Charabati S, Bracco D, Mathieu P, Hemmerling T. Comparison of four different display designs of a novel anaesthetic monitorning system, the Integrated Monitor of Anaesthesia (IMA). Br J Anaesth. 2009; 103(5):670–7. [PubMed: 19767312]
18. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. International Journal of Human Computer Interaction. 1995; 7(1):57–78.
19. Shneiderman, SB.; Plaisant, C. Designing the user interface. 4. Pearson Addison Wesley; USA: 2005.
20. Howell, D. Statistical Methods for Psychology. 5. Pacific Groove, CA: Thomson Learning; 2001.
21. Zhang Y, Drews F, Westenskow D, Foresti S, Agutter J, Bermudez J, et al. Effects of integrated graphical displays on situation awareness in anaesthesiology. Cognit, Technol, & Work. 2002; 4:82–90.
22. Gorges M, Staggers N. Evaluations of physiological monitoring displays: A systematic review. J Clin Monit Comput. 2008; 22:45–66. [PubMed: 18064532]

Research Highlights

- Nurses found the integrated graphical display more usable than their current hospital system.
- Performance improved using the graphical display only at the primary design institution.
- Medical information displays optimized for use at one clinical site may not be as favorable for other clinical sites.
- Health information technology needs to be flexible and configurable to meet local design constraints.

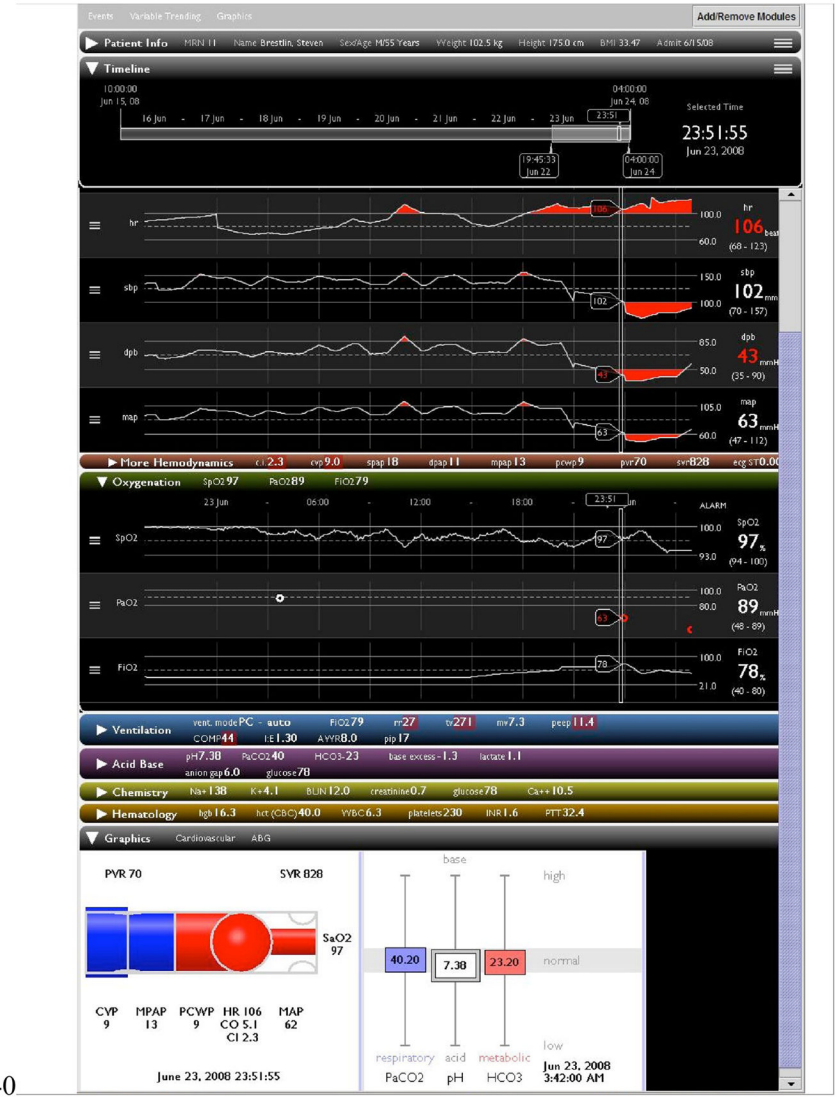
Summary points

What was already known on the topic:

- High-quality ICU patient care depends in part on the rapid and accurate detection of clinical cues revealing possible complications or adverse events.
- Current medical displays are often incompatible with practitioners' workflow and fragment patient information; thus limiting a practitioners' ability to detect evolving changes, and create care inefficiencies and frustrate clinicians.
- New integrated information display systems have the potential to decrease detection times, increase efficiency, improve situation awareness, and lead to better adverse event detection.

What this study added to our knowledge:

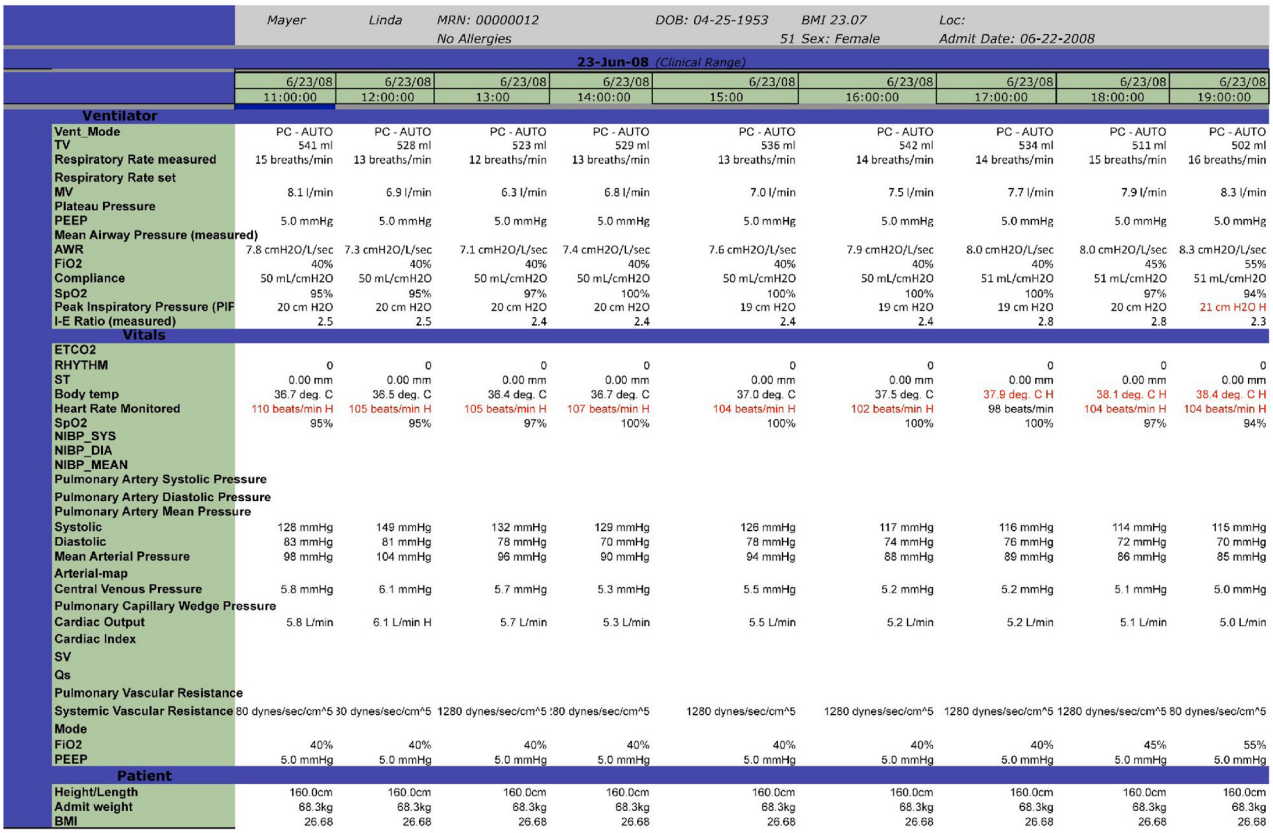
- Nurse expressed greater usability ranking, reported no difference in workload, and at the primary design institution performance improved using an integrated graphical display.
- Medical information displays that have been optimized for use at one clinical site may not be optimal for other clinical sites.
- Health information technology should be flexible and configurable enough to conform to local preferences and conventions, while maintaining any universal components.



251658240

Figure 1.
Graphical display that contains trending and configural display information

251658240



	Mayer	Linda	MRN: 00000012 No Allergies		DOB: 04-25-1953	BMI 23.07 51 Sex: Female	Loc: Admit Date: 06-22-2008		
23-Jun-08 *(Clinical Range)*									
	6/23/08 11:00:00	6/23/08 12:00:00	6/23/08 13:00	6/23/08 14:00:00	6/23/08 15:00	6/23/08 16:00:00	6/23/08 17:00:00	6/23/08 18:00:00	6/23/08 19:00:00
Ventilator									
Vent_Mode	PC - AUTO	PC - AUTO	PC - AUTO	PC - AUTO	PC - AUTO	PC - AUTO	PC - AUTO	PC - AUTO	PC - AUTO
TV	541 ml	528 ml	523 ml	529 ml	536 ml	542 ml	534 ml	511 ml	502 ml
Respiratory Rate measured	15 breaths/min	13 breaths/min	12 breaths/min	13 breaths/min	13 breaths/min	14 breaths/min	14 breaths/min	15 breaths/min	16 breaths/min
Respiratory Rate set									
MV	8.1 l/min	6.9 l/min	6.3 l/min	6.8 l/min	7.0 l/min	7.5 l/min	7.7 l/min	7.9 l/min	8.3 l/min
Plateau Pressure									
PEEP	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg
Mean Airway Pressure (measured)									
AWR	7.8 cmH2O/L/sec	7.3 cmH2O/L/sec	7.1 cmH2O/L/sec	7.4 cmH2O/L/sec	7.6 cmH2O/L/sec	7.9 cmH2O/L/sec	8.0 cmH2O/L/sec	8.0 cmH2O/L/sec	8.3 cmH2O/L/sec
FiO2	40%	40%	40%	40%	40%	40%	40%	45%	55%
Compliance	50 mL/cmH2O	50 mL/cmH2O	50 mL/cmH2O	50 mL/cmH2O	50 mL/cmH2O	50 mL/cmH2O	51 mL/cmH2O	51 mL/cmH2O	51 mL/cmH2O
SpO2	95%	95%	97%	100%	100%	100%	100%	97%	94%
Peak Inspiratory Pressure (PIF	20 cm H2O	20 cm H2O	20 cm H2O	20 cm H2O	19 cm H2O	19 cm H2O	19 cm H2O	20 cm H2O	21 cm H2O H
I-E Ratio (measured)	2.5	2.5	2.4	2.4	2.4	2.4	2.8	2.8	2.3
Vitals									
ETCO2									
RHYTHM	0	0	0	0	0	0	0	0	0
ST	0.00 mm	0.00 mm	0.00 mm	0.00 mm	0.00 mm	0.00 mm	0.00 mm	0.00 mm	0.00 mm
Body temp	36.7 deg. C	36.5 deg. C	36.4 deg. C	36.7 deg. C	37.0 deg. C	37.5 deg. C	37.9 deg. C H	38.1 deg. C H	38.4 deg. C H
Heart Rate Monitored	110 beats/min H	105 beats/min H	105 beats/min H	107 beats/min H	104 beats/min H	102 beats/min H	98 beats/min	104 beats/min H	104 beats/min H
SpO2	95%	95%	97%	100%	100%	100%	100%	97%	94%
NIBP_SYS									
NIBP_DIA									
NIBP_MEAN									
Pulmonary Artery Systolic Pressure									
Pulmonary Artery Diastolic Pressure									
Pulmonary Artery Mean Pressure									
Systolic	128 mmHg	149 mmHg	132 mmHg	129 mmHg	126 mmHg	117 mmHg	116 mmHg	114 mmHg	115 mmHg
Diastolic	83 mmHg	81 mmHg	78 mmHg	70 mmHg	78 mmHg	74 mmHg	76 mmHg	72 mmHg	70 mmHg
Mean Arterial Pressure	98 mmHg	104 mmHg	96 mmHg	90 mmHg	94 mmHg	88 mmHg	89 mmHg	86 mmHg	85 mmHg
Arterial-map									
Central Venous Pressure	5.8 mmHg	6.1 mmHg	5.7 mmHg	5.3 mmHg	5.5 mmHg	5.2 mmHg	5.2 mmHg	5.1 mmHg	5.0 mmHg
Pulmonary Capillary Wedge Pressure									
Cardiac Output	5.8 L/min	6.1 L/min H	5.7 L/min	5.3 L/min	5.5 L/min	5.2 L/min	5.2 L/min	5.1 L/min	5.0 L/min
Cardiac Index									
SV									
Qs									
Pulmonary Vascular Resistance									
Systemic Vascular Resistance	80 dynes/sec/cm^5	30 dynes/sec/cm^5	1280 dynes/sec/cm^5	80 dynes/sec/cm^5	1280 dynes/sec/cm^5	1280 dynes/sec/cm^5	1280 dynes/sec/cm^5	1280 dynes/sec/cm^5	80 dynes/sec/cm^5
Mode									
FiO2	40%	40%	40%	40%	40%	40%	40%	45%	55%
PEEP	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg	5.0 mmHg
Patient									
Height/Length	160.0cm	160.0cm	160.0cm	160.0cm	160.0cm	160.0cm	160.0cm	160.0cm	160.0cm
Admit weight	68.3kg	68.3kg	68.3kg	68.3kg	68.3kg	68.3kg	68.3kg	68.3kg	68.3kg
BMI	26.68	26.68	26.68	26.68	26.68	26.68	26.68	26.68	26.68

Figure 2.
Vital sign information in the tabular display

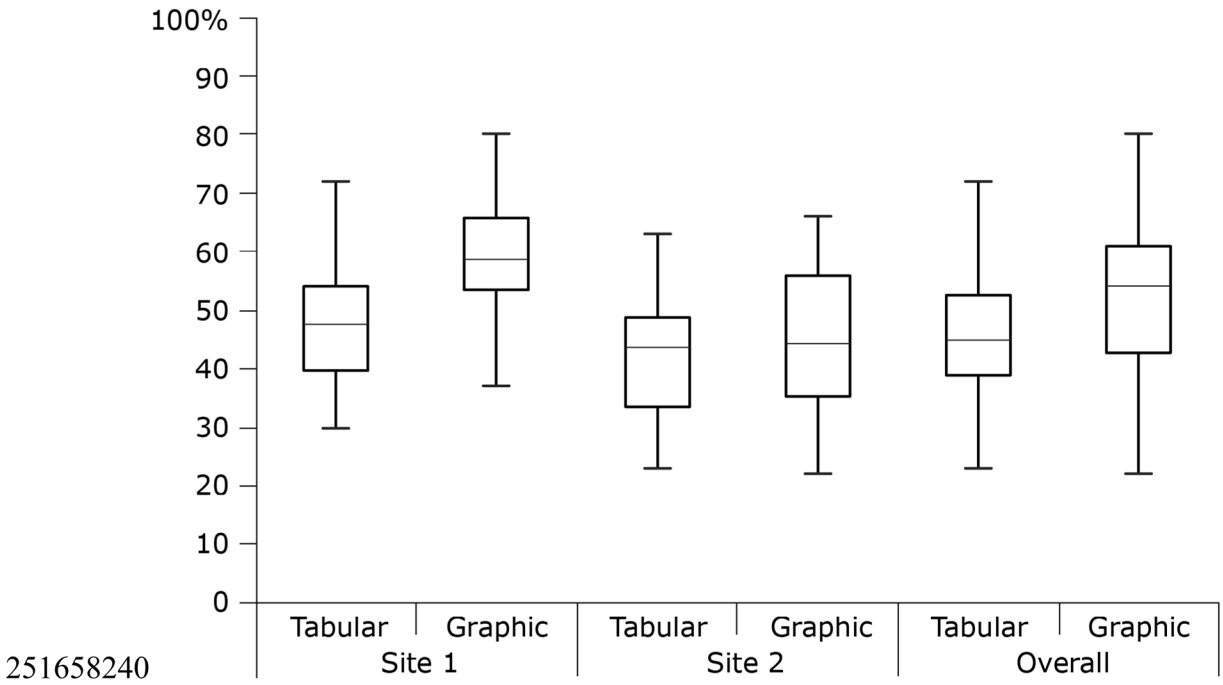


Figure 3.
Boxplots representing median, quartiles, minimum and maximum of Delphi performance score percentage for each location using the tabular and IGID displays (* significant at $p<0.0167$)

251658240

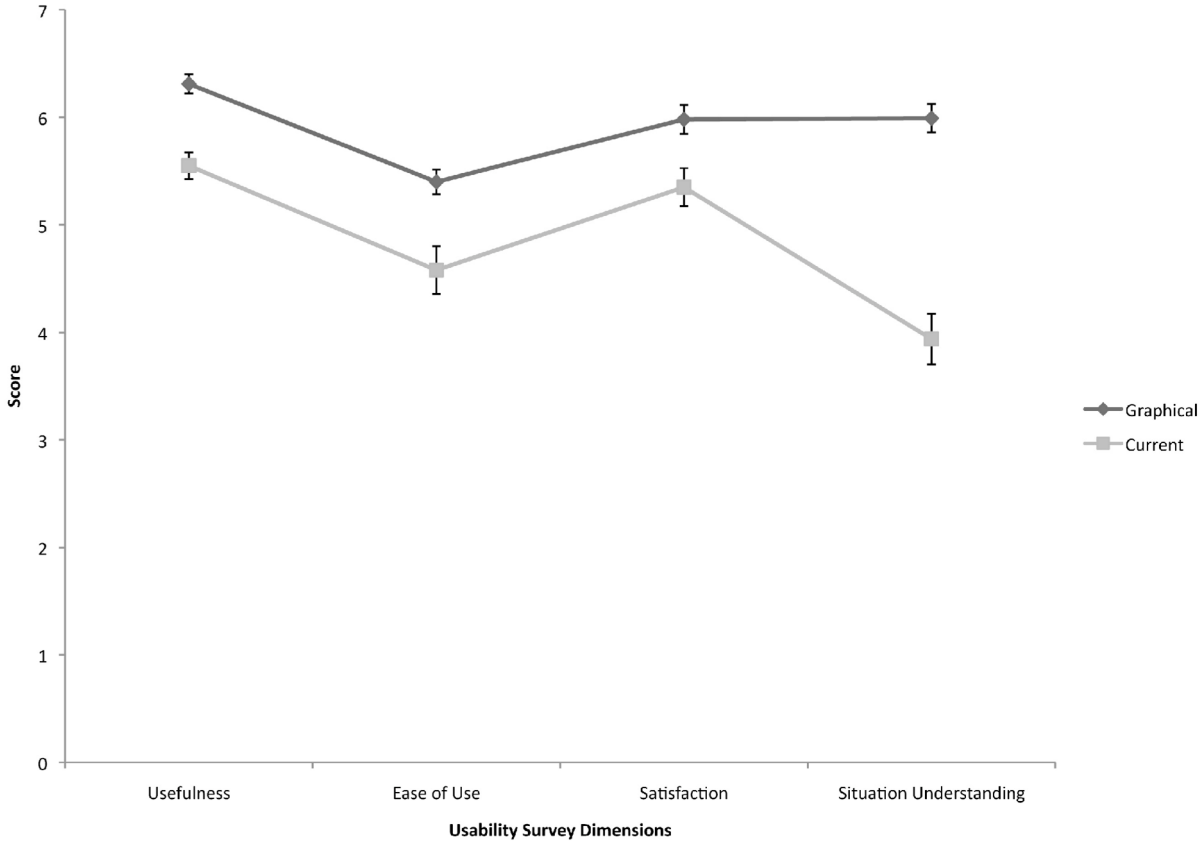


Figure 4.
Mean scores for each dimension of the usability survey for the current in practice display as compared to the IGID display, with mean standard error shown by error bars (all are significant at $p<0.0167$)

Table 1

Study site profile

	Site A	Site B
General profile	University teaching hospital	Major metropolitan and regional not-for profit university teaching hospital
Geographical location	North-western United States	Southern-central United States
Number of ICU beds	Surgical 20 Medical 12 Burn 12 Neurological 18	Surgical 21 Medical 26 Trauma 31 Burn 25 Neurological 22 Cardio-vascular 27
Average length of patient stay (in days)	Surgical: 3.68 Medical: 3.1 Burn: 12.5 Neurological: 3.4	Surgical: 5.1 Medical: 4.3 Trauma: 6.1 Burn: 5.7 Neurological: 5.1 Cardio-vascular: 4.3
Average number of nurses a day per unit	Surgical: 20 Medical: 12 Burn: 12 Neurological: 20	Surgical: 17 Medical: 19 Trauma: 16 Burn: 11 Neurological: 14 Cardio-vascular: 16
Type of ICU medical decision making	Surgical: open with intensivist led ICU team Medical: closed, intensivist led Burn: closed Neurological: open, partial intensivist coverage	Surgical: closed Medical: closed Trauma: closed Burn: closed Neurological: closed Cardio-Vascular: open, partial intensivist coverage
Local clinician involvement in UI design & development	Iterative UCD involvement, Delphi scoring of evaluation criteria	Limited to one round of Delphi scoring only

Table 2

Usability survey statements and sub-groupings *

Grouping	Statement
Usefulness	The display contains information that would be useful to me. ¢ The display supports my decision-making. ¢ Standard medical language is used throughout the display. ¢ The display presents information in a meaningful way. ‡ I use the information on the display to inform my thinking about the patient's situation. ¢
Ease of Use	The display is easy to understand.‡ The organization of the information on the display makes sense to me.‡ 90% of the people I work with can use this display the first time they try it. I am frustrated when using this display.‡ With training, most of the people I work with are able to use this display effectively. ¢
Satisfaction	I resist using this display in the ICU. I recommend that other people in my unit use this display. I would use the new display in addition to my current equipment. This display is flexible enough to help me do my job.‡ I think this display is an improvement on the current display. I am enthusiastic about using this display.‡
Situation	I can see drugs in the context of other patient variables.
Understanding	When using the display, I am able to develop a complete picture of the patient's situation. The relationship between cause and effect is apparent when using the display.
Graphics	The graphic representation of trended data is useful. Seeing the variables together on a timeline helps me understand what is occurring. The ABG graphic improves my understanding of the patient's state. The cardiovascular graphic improves my understanding of the patient's state.

* Adapted from common usability questionnaires (18, 19) with additional questions focused on situation understanding and display specific elements.

Table 3

Possible reasons for the apparent site differences

Possible Reasons for Site Differences	Comments
Sampling bias.	Participants were motivated volunteers who were paid for their time. The samples may not have been representative of the populations of all nurses at the sites.
Differences in conduct of the study.	Less likely since both sites adhered to a script and study conduct uniformity was confirmed during scoring.
Differences in application of scoring criteria.	Unlikely since all data were scored collectively in a uniform manner by the research team.
Differences in the typical electronic chart (display) used at the two sites.	The two sites used different commercial ICU EHR systems. There may have been more positive (or negative) transfer of training at one site to either of the displays. However, the graphic display was novel and the tabular display was an intentional hybrid of the two commercial systems.
Differences in participant background or demographics.	Site B nurses had somewhat more clinical experience but this was included as a covariate in the statistical analysis.
Differences in nurses' clinical practices, customs, and preferences.	Site B nurses had been trained in the use of SBAR-based structured event reporting to physicians.
Use of Site A as primary source of clinician input during graphic display design.	The graphic display design may have been over-optimized for the display use practices and preferences of ICU nurse users at Site A.