Comparing two anesthesia information management system user interfaces: a usability evaluation

Comparaison de l'interface utilisateur de deux systèmes de gestion de l'information pour l'anesthésie: évaluation de leur facilité d'emploi

Jonathan P. Wanderer, MD · Anoop V. Rao, MBBS ·
Sarah H. Rothwell, MHS · Jesse M. Ehrenfeld, MD

Abstract

Introduction *Anesthesia information management systems (AIMS) have been developed by multiple vendors and are deployed in thousands of operating rooms around the world, yet not much is known about measuring and improving AIMS usability. We developed a methodology for evaluating AIMS usability in a low-fidelity simulated clinical environment and used it to compare an existing user interface with a revised version. We hypothesized that the revised user interface would be more useable.*

Methods *In a low-fidelity simulated clinical environment, twenty anesthesia providers documented essential*

J. P. Wanderer, MD (✉) · J. M. Ehrenfeld, MD
Vanderbilt University, 1301 Medical Center Drive, TVC 4648,
Nashville, TN 37212, USA
e-mail: Jonathan.p.wanderer@vanderbilt.edu

A. V. Rao, MBBS
Department of Anesthesia, Critical Care and Pain, Massachusetts
General Hospital, Gray-Bigelow 444 55 Fruit Street, Boston,
MA 02114, USA

S. H. Rothwell, MHS
Biomedical Engineering, Massachusetts General Hospital,
Boston, MA, USA

anesthetic information for the start of the case using both an existing and a revised user interface. Participants had not used the revised user interface previously and completed a brief training exercise prior to the study task. All participants completed a workload assessment and a satisfaction survey. All sessions were recorded. Multiple usability metrics were measured. The primary outcome was documentation accuracy. Secondary outcomes were perceived workload, number of documentation steps, number of user interactions, and documentation time. The interfaces were compared and design problems were identified by analyzing recorded sessions and survey results.

Results *Use of the revised user interface was shown to improve documentation accuracy from 85.1% to 92.4%, a difference of 7.3% (95% confidence interval [CI] for the difference 1.8 to 12.7). The revised user interface decreased the number of user interactions by 6.5 for intravenous documentation (95% CI 2.9 to 10.1) and by 16.1 for airway documentation (95% CI 11.1 to 21.1). The revised user interface required 3.8 fewer documentation steps (95% CI 2.3 to 5.4). Airway documentation time was reduced by 30.5 seconds with the revised workflow (95% CI 8.5 to 52.4). There were no significant time differences noted in intravenous documentation or in total task time. No difference in perceived workload was found between the user interfaces. Two user interface design problems were identified in the revised user interface.*

Discussion *The usability of anesthesia information management systems can be evaluated using a low-fidelity simulated clinical environment. User testing of the revised user interface showed improvement in some usability metrics and highlighted areas for further revision. Vendors of AIMS and those who use them should consider adopting methods to evaluate and improve AIMS usability.*

Résumé

Introduction *Les systèmes de gestion de l'information pour l'anesthésie (SGIA) ont été développés par de nombreux fournisseurs et sont déployés dans des milliers de blocs opératoires dans le monde et, pourtant, on sait peu de choses sur les façons de mesurer et d'améliorer la facilité d'emploi de ces SGIA. Nous avons élaboré une méthodologie pour l'évaluation de la facilité d'emploi des SGIA dans un environnement clinique simulé à basse fidélité et nous l'avons utilisée pour établir une comparaison avec une version révisée d'une interface utilisateur existante. Nous avons émis l'hypothèse que l'interface utilisateur révisée serait plus facile d'utilisation.*

Méthodes *Dans un environnement clinique simulé à basse fidélité, vingt professionnels dans le domaine de l'anesthésie ont documenté l'information essentielle sur l'anesthésie pour le début d'un cas en utilisant à la fois une interface utilisateur existante et une interface révisée. Les participants n'avaient pas utilisé la version révisée précédemment et ils ont suivi un court exercice de formation avant d'accomplir la tâche prévue par l'étude. Tous les participants ont répondu à une enquête sur l'évaluation de la charge de travail et leur satisfaction. Toutes les sessions ont été enregistrées. De nombreux critères de facilité d'emploi ont été mesurés. Le critère d'évaluation principal était l'exactitude de la documentation. Les critères d'évaluation secondaires étaient la charge de travail perçue, le nombre d'étapes dans la documentation, le nombre d'interactions avec l'utilisateur et le temps de documentation. Les interfaces ont été comparées et les problèmes de conception ont été identifiés à l'aide de l'analyse des séances enregistrées et des résultats de l'enquête.*

Résultats *Il a été montré que l'utilisation d'une interface utilisateur révisée améliorait l'exactitude de la documentation de 85,1 % à 92,4 %, une différence de 7,3 % (intervalle de confiance [IC] à 95 % pour la différence: 1,8 à 12,7). Avec l'interface révisée, le nombre d'interactions de l'utilisateur a été réduit de 6,5 (IC à 95 %: 2,9 à 10,1) pour la documentation intraveineuse et de 16,1 pour la documentation sur voies aériennes (IC à 95 %: 11,1 à 21,1). L'interface utilisateur a nécessité 3,8 étapes de documentation de moins (IC à 95 %: 2,3 à 5,4). La documentation sur voies aériennes a pris 30,5 secondes de moins avec le flux de travail révisé (IC à 95 %: 8,5 à 52,4). Il n'y a pas eu de différences significatives en termes de temps pour la documentation intraveineuse ou pour la durée totale de la tâche. Aucune différence n'a été observée concernant la charge de travail perçue entre les deux interfaces utilisateur. Deux problèmes de conception de l'interface utilisateur ont été identifiés dans l'interface révisée.*

Discussion *La facilité d'emploi des systèmes de gestion de l'information pour l'anesthésie peut être évaluée en utilisant un environnement clinique simulé à basse fidélité. Les tests par des utilisateurs de l'interface révisée ont montré une amélioration de l'utilisation de certains outils de mesure et ont fait apparaître des domaines nécessitant des révisions complémentaires. Les distributeurs de SGIA et ceux qui les utilisent doivent envisager l'adoption de méthodes destinées à évaluer et à améliorer la facilité d'emploi des SGIA.*

Introduction

The adoption of anesthesia information management systems (AIMS) has the potential to improve patient care by facilitating support for real-time decisions, promoting clinician communication, and increasing revenue through timely billing.[1-5] However, these benefits are dependent on the ability of clinicians to use AIMS effectively and to enter data in an accurate and timely manner.[6,7] Software usability is defined as "the extent to which a product can be used by specific users to achieve specific goals with effectiveness, efficiency and satisfaction in the specified context of use."[8] We define AIMS usability as the extent to which an anesthesia provider is effective, efficient, and satisfied with documenting an anesthetic.

While AIMS have been developed by multiple vendors and deployed in operating rooms around the world, little is known about measuring and improving AIMS usability. Data from one survey of anesthesiologists using AIMS show that 29% had less efficient workflow; 49% had an unanticipated need for ongoing information technology support; 19% had inaccurate records, and only 43% were very satisfied with their AIMS implementation.[9] Developing a methodology for evaluating AIMS usability might be helpful to determine the extent to which these reported problems are related to user interface design and also to facilitate comparison of AIMS products.

The AIMS used in our institution permits extensive user interface customization, and we have developed a revision of our existing user interface to improve our AIMS usability. The purpose of this study was to develop a methodology for measuring AIMS usability and to assess our revised AIMS user interface against the existing user interface. Evaluating our AIMS user interface necessitated drawing on user interaction methodology from other fields where such processes are formalized and used extensively. We used an evaluation strategy in which study subjects charted anesthetic inductions in a low-fidelity simulated clinical environment while multiple usability metrics were measured. We selected documentation accuracy as the primary outcome. Secondary outcomes included three separate efficiency metrics: length of time required to

complete tasks, number of user interactions (key presses and screen touches) needed to complete specific subtasks, and total number of documentation steps. Documentation accuracy was the effectiveness metric. Two additional secondary outcomes were measured to assess user satisfaction: a validated workload assessment tool and an *ad hoc* survey. Using this methodology, we hypothesized that the revised user interface would be more usable than the existing user interface.

Methods

Participants

The study protocol was approved by the Partners Human Research Committee, and the requirement for written informed consent was waived. Twenty voluntary participants (ten residents and ten staff anesthesiologists) were tested individually in single sessions that lasted approximately 45 min over a two-week time period. A power calculation was not performed *a priori*. A sample size of 20 was chosen based on prior work[10-16] due to a lack of an estimate of measured effect size.

Existing user interface

The AIMS used in our institution is MetaVision (iMDsoft®, Needham, MA, USA). Like other AIMS, it required initial customization to support existing clinical workflow. In the existing user interface, case event buttons are arranged linearly from left to right on a toolbar at the top of the screen (Fig. a), and submenus can be opened for additional events (Fig. b). For ease of documentation, frequently used documentation elements are represented in more than one place (e.g., the "Lines" button is on the main toolbar and also in the submenu, Fig. a). New clinicians receive a one-hour orientation to our AIMS from a support team member, and they receive additional instruction by shadowing current clinicians. Clinicians use this interface for all anesthetics administered.

Revised user interface

The revised user interface that we tested (Fig. c, d) differed from our existing user interface in the following ways: 1) The documentation process provides continuous visual feedback on completeness. For example, successful completion of vascular access documentation is indicated by the "Access" tab changing from red to green. Once all of the task tabs are green, an indicator in the main chart changes from red to green, indicating that all crucial documentation is complete. Red and green colours with contrasting gray scale properties were selected and tested in order to accommodate colour-blind individuals. 2) Documentation elements are context-sensitive. While there are a number of potentially critical aspects to airway documentation, not all are relevant to a given airway approach. Documentation of direct laryngoscopy, for instance, automatically prompts for Cormack-Lehane grade and remains hidden if only a mask airway is documented. 3) Documentation workflow is matched to typical clinical workflow. For instance, documentation of the time of intubation was separated from intubation details to allow timely documentation. 4) The documentation process is designed to minimize AIMS user interactions. Frequently used drop-down lists were paired with buttons that allowed a single-click selection of default values based on previously documented patient characteristics. Study participants had never utilized the user interface prior to this study.

Apparatus

Participants charted anesthetic inductions in a low-fidelity simulated clinical environment equipped with an anesthesia machine and our institution's standard AIMS interface. The environment was set up in a manner identical to our institution's general operating room with a monitor at eye level attached to the anesthesia machine. An additional monitor displayed patient scenarios and the training video (available as Electronic Supplemental Material) and permitted completion of an electronic questionnaire. The environment did not include a mannequin, actors, drugs, or anesthetic equipment aside from the anesthesia machine and AIMS interface. Participants used a touch screen and keyboard to document. A screen recording program was used to capture participants' AIMS sessions, (CamStudio, RenderSoft, Emeryville, CA, USA), and a video camera was used to record participants' interactions with the AIMS software.

Experimental design

Participants took part on an "as available" basis and were sequentially assigned to two groups. Using a counterbalanced design, the first group performed tasks with the existing user interface followed by the revised user interface. The second group performed tasks with the revised user interface followed by the existing user interface. The results from both groups were combined for analysis. Tasks were balanced such that the number of documentation elements in each case was identical. Cases were balanced such that each was documented an equal number of times with each user interface.

Figure a) Existing user interface with case events arranged linearly from left to right on toolbar and initial menu opened. b) Intravenous documentation submenu opened in existing user interface c) Intravenous documentation in revised user interface d) Context-sensitive airway documentation in the revised user interface. Task buttons change color from red to green to provide continuous visual feedback on completeness
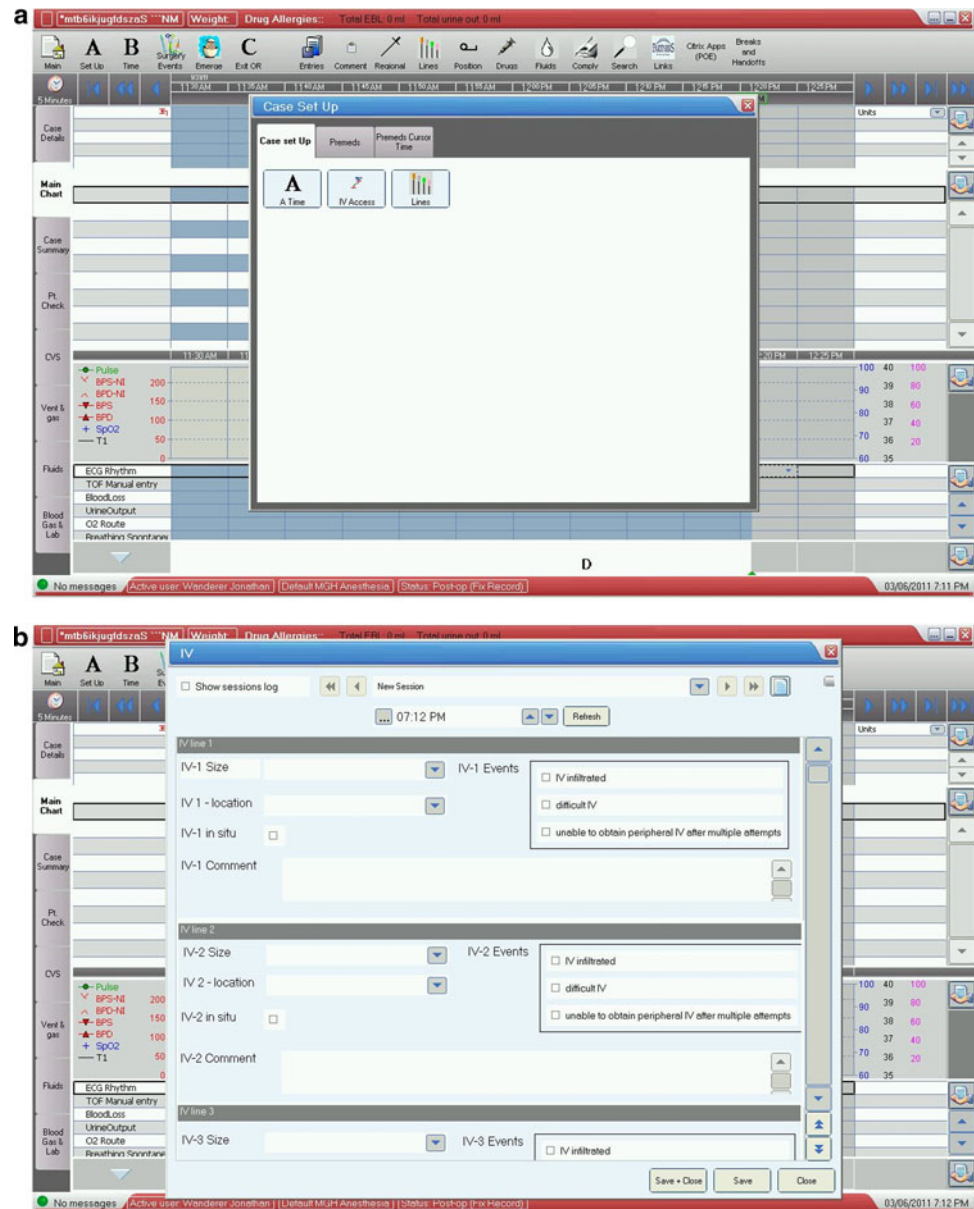
Procedure

Participants completed a questionnaire soliciting experience level, age, frequency of assuming primary responsibility for documentation of an anesthetic, and length of time using the existing user interface of our AIMS. Each group then documented the beginning of three separate cases (available as Electronic Supplemental Material). The participants in the first group documented the start of the first case using the existing user interface, and they then documented the start of the second case using the revised user interface while watching a training video. The first group finished by documenting the start of the third case using the revised user interface without assistance. The participants in the second group started with the revised user interface and training video, followed by documentation using the revised user interface. They finished by documenting the start of the last case using the existing user interface. The training sessions were not analyzed. After each of the two experimental documentation tasks, participants rated their perceived workload by completing the National Aeronautical Space Administration Task Load Index (NASA-TLX) worksheet (available as Electronic Supplemental Material). After the end of the last task, participants completed an *ad hoc* survey assessing satisfaction with the individual elements that were revised and provided feedback (available as Electronic Supplemental Material).

Figure continued

Measurement of interactions, tasks, subtasks and documentation steps

The screen recordings from each participant session were reviewed. The overall task was divided into five subtasks corresponding with the revised user interface tabs: vascular access, drugs, airway, positioning, and patient checks (Fig. c). Overall task time and subtask time were determined by measuring the interval between the first and last interaction required to start and finish the task, respectively. Subtask analysis was restricted to vascular access and airway, as the user interface for the drugs, positioning, and patient checks were similar in both the revised and existing user interface. A user interaction was defined as either a key press or finger press on the touch screen. The number of documentation steps was determined by counting the number of menus utilized to complete a task, including redundantly accessed menus.

Measurement of documentation accuracy

Anesthetic records were analyzed for accuracy and completeness by comparison with an anesthetic record in which every specified element (available as Electronic Supplemental Material) was recorded accurately. Accuracy of each task was determined by the percentage of correct data elements associated with each task. Missing and erroneous data elements were treated equally as incorrect.

Data analysis

Data were analyzed using SPSS Statistics version 17.0 (SPSS, Inc., Chicago, IL, USA). Outcomes were paired and calculated as the difference of the means with 95% confidence intervals (CI) for the difference. Statistical significance was determined by lack of overlap between the 95% CI and zero. Correlations between dependent values were assessed using the Pearson method. Interactions between and within groups were evaluated as independent variables using linear regression with age, academic rank, AIMS usage frequency, length of AIMS usage, and presentation order.

Results

Twenty participants (10 clinical anesthesia (CA) residents - six CA-1, one CA-2, three CA-3, and ten attending anesthesiologists) completed the study protocol. However, a technical error in a screen recording led to the exclusion of one attending participant from analysis. Nine of the resident participants reported performing induction documentation at least once a day, with the remaining resident performing the task at least once a week. Six of the attending participants reported performing induction documentation at least once a day; two performed the task at least once a week, and two performed the task at least once a month. Average length of time (standard deviation) using AIMS with the existing user interface was 22.2 (16.7) mth (range 1-8 mth). All participants had used the existing user interface for at least one month prior to participating.

Analysis of effectiveness

The primary outcome was documentation accuracy. Thirty documentation elements were specified in each case scenario (available as Electronic Supplemental Material). The head support documentation element was removed from analysis due to a design problem that interfered with its display in the revised user interface. Documentation accuracy with the revised user interface was 92.4% compared with an accuracy of 85.1% with the existing user interface. The mean difference was 7.3% (95% CI for the difference 1.8 to 12.7).

Measuring workload and satisfaction

The Table lists the NASA TLX composite score for each user interface. No statistically significant differences in workload were noted. The satisfaction survey (see Electronic Supplemental Material) did not directly compare the revised and existing user interfaces and did not assess for satisfaction with specific elements of the existing user interface. Participants agreed with the statement, "Overall the new case start form was easy to use" (median 6; range 4-7 on Likert scale with 1 representing disagree strongly, 7 representing agree strongly). They agreed with the statement, "This case start form would work well within my existing workflow" (median 6; range 3-7), and they agreed with the statement, "I would like to use this case start form in my OR" (median 6; range 3-7). Users indicated satisfaction with the revised elements, including the use of colour-coded buttons (median 6; range 4-7) and graphical icons (median 6; range 2-7).

Assessing efficiency

The number of user interactions and time required to complete documentation subtasks are listed in the Table. Use of the revised user interface showed a statistically significant decrease in the number of user interactions as well as a decrease in the time required to complete airway documentation. Intravenous documentation using the revised user

Table Aims Usability Metrics

	Existing user interface mean (95% CI)	Revised user interface mean (95% CI)	Difference mean (95% CI)
Total Task Time (seconds)	208.7 (178.9 to 238.5)	243.8 (213.3, 274.4)	35.1 (-0.9 to 69.2)
NASA TLX	13.2 (10.5 to 15.9)	12.6 (9.5 to 15.7)	-0.6 (-3.4 to 2.1)
Documentation Accuracy (% correct)	85.1 (77.9 to 92.2)	92.4 (89.3 to 95.4)	7.3 (1.8 to 12.7)
Documentation Steps (n)	10.7 (8.9 to 12.5)	6.8 (6.4 to 7.3)	-3.8 (-2.3 to -5.4)
Intravenous Task Time (seconds)	30.8 (26.7 to 35.0)	29.7 (20.3 to 39.0)	-1.2 (-12.0 to 9.7)
Intravenous Task User Interactions (n)	15.6 (13.6 to 17.7)	9.1 (6.0 to 12.3)	-6.5 (-2.9 to -10.1)
Airway Task Time (seconds)	75.1 (51.1 to 99.0)	44.6 (39.1 to 50.2)	-30.5 (-8.5 to -52.4)
Airway Task User Interactions (n)	32.2 (27.5 to 36.9)	16.1 (14.6 to 17.6)	-16.1 (-11.1 to -21.1)

CI = confidence interval; AIMS = anesthesia information management systems; NASA TLX = National Aeronautical Space Administration task load index - six categories are scored 1-7 on a Likert scale, shown is composite score. Documentation steps were defined as the number of menus accessed to complete the overall task. User interactions are defined as the number of key strokes plus the number of screen touches used to complete the task

interface showed a reduction in the number of user interactions needed with no difference in time required. The revised user interface was associated with substantially fewer documentation steps. The majority of participants using the revised user interface did so with the optimal number of documentation steps, which we defined as a unidirectional linear traversal of the submenus. In contrast, with the existing user interface, no two participants performed documentation in the same manner. Total documentation time did not show a statistically significant difference between the two user interfaces although there was a trend towards increased documentation time with the revised user interface.

Identification of design problem

The revised user interface was developed in response to the perceived design problems of the existing user interface, specifically the lack of feedback on documentation completeness, high levels of required user interaction, and mismatch between our AIMS and clinical workflows. Thus, the existing user interface was not subjected to additional review for design problems. Review of recorded sessions and survey feedback revealed two design problems in the revised user interface that had not been identified previously. First, as mentioned above, the head support documentation element was not consistently visible to the user. This likely led to increased time spent searching for this documentation element. Second, several users made incorrect entries in documenting intravenous placement and spent substantial time attempting to correct these entries. While entries could be modified, the mechanism for removing an erroneous entry was not intuitive. Both of these design issues increased the total task time.

Measurement interactions

A significant group effect was found between AIMS usage frequency and documentation steps with the existing user interface ($\beta = 6.17$; 95% CI 2.65 to 9.70; $P = 0.002$), indicating that frequent users were more efficient at documentation. No other interactions were noted. Testing for correlated dependent variables revealed multiple correlations between task time and task user interactions as well as correlations within task times, all expected findings as each dependent variable was chosen to measure usability.

Discussion

The revised user interface led to improved documentation accuracy, fewer user interactions, fewer documentation steps, and decreased time required to chart airway details.

Furthermore, many users found the revised user interface easy to use despite not having used it previously. Significant differences in cognitive workload were not observed, potentially because the tasks being compared did not present a substantial cognitive challenge. We also did not find a difference in overall task completion time, which may be related to lack of experience with the revised user interface. Overall, we have shown that it is possible to revise an AIMS user interface and show improvements in usability using a low-fidelity simulated clinical environment.

There are a number of alternatives to the evaluation of AIMS user interfaces in low-fidelity simulated clinical environments. For instance, high-fidelity simulated clinical environments have been used to evaluate the design of specific anesthesia delivery systems[16,17] and to examine interaction with certain aspects of the system's design, such as pressure limitation in volume control ventilation during the onset of bronchospasm. Creating such a specific scenario necessitates a high-fidelity environment so that study participants can flip physical switches and perform auscultation to reach correct clinical diagnoses. In contrast, evaluation of AIMS user interfaces can be performed with fewer resources by removing non-essential elements from the simulation environment, such as mannequins and surgical actors. Another approach to evaluating aspects of the anesthesia environment is observation of actual clinical activity with behavioural task analysis. Like high-fidelity simulation, this method is more resource intensive and requires rigorous personnel training.[15] Additionally, evaluation of new technology in actual clinical environments is potentially less safe than doing so in simulated conditions. Evaluation of AIMS user interfaces in low-fidelity simulated clinical environments lowers the bar for usability testing and facilitates clinician input into AIMS design and revisions early in the process.

As recent national initiatives from the Agency of Healthcare Research and Quality,[A] the Health Information Management Systems Society,[8] and the National Institute for Standards and Technology[18] show, usability remains an important concern as the adoption of AIMS by anesthesia practices continues and hospitals implement electronic health records. In addition to increasing effectiveness, efficiency, and user satisfaction, improvements in usability have also been shown to decrease costs associated with software training, support, and maintenance.[19] In contrast, poor usability has been associated with decreased productivity and user frustration.[20,21] In a national physician survey that examined electronic health record adoption

barriers, usability concerns ranked second and third behind only cost.[22]

The specific AIMS package used at our institution can be customized easily, which represents a double-edged sword for system designers and clinicians. The potential for creating usable user interfaces that precisely match clinical workflow is balanced by the ability to construct poor user interfaces that interfere with patient care. Other AIMS manufacturers adopt an alternative approach whereby they offer user interfaces that are more difficult to configure, i.e., there is reduced workflow flexibility and less potential to deviate from the manufacturers' design. While currently it remains unclear as to which approach is preferable, usability testing offers a method to compare within and between AIMS offerings, and the adoption of usability testing will facilitate incremental improvements.

Optimal user interface design remains a challenge in the medical field where there is little overlap between designers and users and where successful design requires in-depth understanding of human factor issues.[23,24] The focus of prior work related to AIMS human factors has been on comparing paper-based documentation with electronic documentation rather than on comparing two electronic solutions.[10] One of the important aspects of this study was the use of a multimodal approach in user interface evaluation that allowed for the analysis of specific aspects of the design. Review of screen recordings provided an objective evaluation of specific changes, while subjective input was obtained from the *ad hoc* survey. Combining these data facilitates formulating multiple simultaneous alterations to a user interface while retaining the ability to evaluate modifications individually, thus potentially leading to rapid improvement.

A key feature of the simulation environment was the recording of sessions via both screen and video recording. This approach allowed for in-depth interaction analysis by viewing both recordings together. This was particularly helpful in identifying user interface problems. For instance, the lack of an obvious mechanism for correcting erroneous intravenous documentation was readily apparent on review. Additionally, interaction analysis showed that the revised user interface's colour-coded completeness feedback was effective in encouraging participants to return to prior menus to address areas of incomplete data. These insights were not evident in informal user testing and have guided further user interface revisions.

Measuring user interactions for a given task also proved to be a useful metric. Analysis of the revised user interface indicated that replacement of drop-down option boxes with single-click buttons was successful in reducing user interactions. Carefully designed interfaces that give context-sensitive default options with respect to age and patient size may improve the user experience. Anesthesia

information management systems have already been used to improve postoperative nausea and vomiting prophylaxis,[3] decrease gaps in blood pressure monitoring,[5] and improve adherence with perioperative antibiotic administration.[4] While some AIMS implementations are simply electronic versions of paper charts, AIMS hold the potential to be tools that support clinicians in rendering optimal patient care.

Limitations

There are a number of limitations associated with this study. This study focused on a single AIMS implementation at one institution. Not all AIMS implementations allow the modifications described above, and differences in institutional workflow may limit the ability to generalize our findings. Second, while realistic clinical scenarios were provided to clinicians in an environment similar to an anesthetizing area, these conditions still differ substantially from the clinical reality of inducing anesthesia and positioning and starting a procedure. While simulation in a low-fidelity environment such as ours reduces the barriers to performing an evaluation, it also limits the external validity of any conclusions reached as the workflow is incomplete with respect to both a high-fidelity simulation and an actual clinical environment. Third, the events being documented normally occur in a serial but discrete fashion rather than as a continuous stream of events. Fourth, participants had more experience with the existing user interface than with the revised user interface, which may have resulted in improved usability metrics for the existing user interface that would not have occurred with naïve participants. Task time would likely decrease with additional experience with the revised user interface because users perform the same task repeatedly even as the number of user interactions remains constant. As an additional limitation, the perceived difficulty of the task was low, making it challenging to detect reductions in cognitive workload. Regarding the user interactions metric, measurement of the total number of keystrokes plus the total number of button presses does not always accurately represent the underlying complexity of the task, as typing skills vary across users. In addition, typing a long phrase in one text field may be considerably easier than entering the same number of keystrokes across multiple text fields, but this difference would not be reflected in our user interactions metric. Finally, in this study, we did not evaluate either of our AIMS user interfaces against the non-AIMS alternative, the handwritten record, and the methodology described does not lend itself to such a comparison.

In conclusion, anesthesia information management systems user interfaces can be compared successfully using a low-fidelity simulated clinical environment. The revised

user interface resulted in an improvement in documentation accuracy, reduced user interactions, decreased documentation steps, and decreased time required to chart airway details. Additionally, our usability methodology identified areas in need of further modification, and such modifications identified by usability testing can be implemented prior to wide-scale roll-out. This may potentially accelerate interface improvement while improving efficiency and user satisfaction. The described evaluation methodology permits head-to-head comparisons of AIMS user interfaces and facilitates iterative interface design based on user feedback, AIMS records, and review of AIMS screen recording and user interactions. We suggest that designers of AIMS perform usability analysis, such as described in this manuscript, to evaluate and improve AIMS user interfaces.

References

1. *Spring SF*, *Sandberg WS*, *Anupama S*, *Walsh JL*, *Driscoll WD*, *Raines DE*. Automated documentation error detection and notification improves anesthesia billing performance. Anesthesiology 2007; 106: 157-63.

2. *Kheterpal S*, *Gupta R*, *Blum JM*, *Tremper KK*, *O'Reilly M*, *Kazanjian PE*. Electronic reminders improve procedure documentation compliance and professional fee reimbursement. Anesth Analg 2007; 104: 592-7.

3. *Kooij FO*, *Klok T*, *Hollmann MW*, *Kal JE*. Decision support increases guideline adherence for prescribing postoperative nausea and vomiting prophylaxis. Anesth Analg 2008; 106: 893-8.

4. *Nair BG*, *Newman SF*, *Peterson GN*, *Wu WY*, *Schwid HA*. Feedback mechanisms including real-time electronic alerts to achieve near 100% timely prophylactic antibiotic administration in surgical cases. Anesth Analg 2010; 111: 1293-300.

5. *Ehrenfeld JM*, *Epstein RH*, *Bader S*, *Kheterpal S*, *Sandberg WS*. Automatic notifications mediated by anesthesia information management systems reduce the frequency of prolonged gaps in blood pressure documentation. Anesth Analg 2011; 113: 356-63.

6. *Sandberg WS*, *Sandberg EH*, *Seim AR*, et al. Real-time checking of electronic anesthesia records for documentation errors and automatically text messaging clinicians improves quality of documentation. Anesth Analg 2008; 106: 192-201.

7. *Epstein RH*, *Dexter F*, *Ehrenfeld JM*, *Sandberg WS*. Implications of event entry latency on anesthesia information management decision support systems. Anesth Analg 2009; 108: 941-7.

8. *Healthcare Information and Management Systems Society (HIMSS)*. Defining and Testing EMR Usability: Principles and Proposed Methods of EMR Usability Evaluation and Rating: Healthcare Information Management and Systems Society Electronic Health Record Usability Task Force, June 2009. Available from URLhttp://www.himss.org/content/files/himss_definingandtestingemrusability.pdf (accessed April 2012).

9. *Trentman TL*, *Mueller JT*, *Ruskin KJ*, *Noble BN*, *Doyle CA*. Adoption of anesthesia information management systems by US anesthesiologists. J Clin Monit Comput 2011; 25: 129-35.

10. *Chase CR*, *Ashikaga T*, *Mazuzan JE Jr*. Measurement of user performance and attitudes assists the initial design of a computer user display and orientation method. J Clin Monit 1994; 10: 251-63.

11. *Albert RW*, *Agutter JA*, *Syroid ND*, *Johnson KB*, *Loeb RG*, *Westenskow DR*. A simulation-based evaluation of a graphic cardiovascular display. Anesth Analg 2007; 105: 1303-11.

12. *Syroid ND*, *Agutter J*, *Drews FA*, et al. Development and evaluation of a graphical anesthesia drug display. Anesthesiology 2002; 96: 565-75.

13. *Liu Y*, *Osvalder AL*. Usability evaluation of a GUI prototype for a ventilator machine. J Clin Monit Comput 2004; 18: 365-72.

14. *Charabati S*, *Bracco D*, *Mathieu PA*, *Hemmerling TM*. Comparison of four different display designs of a novel anaesthetic monitoring system, the 'integrated monitor of anaesthesia (IMA)'. Br J Anaesth 2009; 103: 670-7.

15. *Slagle J*, *Weinger MB*, *Dinh MT*, *Brumer VV*, *Williams K*. Assessment of the intrarater and interrater reliability of an established clinical task analysis methodology. Anesthesiology 2002; 96: 1129-39.

16. *Dalley P*, *Robinson B*, *Weller J*, *Caldwell C*. The use of high-fidelity human patient simulation and the introduction of new anesthesia delivery systems. Anesth Analg 2004; 99: 1737-41.

17. *Sowb YA*, *Howard SK*, *Raemer DB*, *Feinstein D*, *Fish KJ*, *Gaba DM*. Clinicians' recognition of the Ohmeda Modulus II plus and Ohmeda Excel 210 SE anesthesia machine system mode and function. Simul Healthc 2006; 1: 26-31.

18. *Schumacher RM*, *Lowry SZ*. NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records NIST Interagency Report 7741. Gaithersburg, MD.: National Institute for Standards and Technology, 2010. Available from URL:http://www.nist.gov/customcf/get_pdf.cfm?pub_id=907313 (accessed April 2012).

19. *Bias R*, *Mayhew D*. Cost-Justifying Usability - An Update for the Internet Age. 2nd ed. San Francisco: Morgan Kaufmann Publishers; 2005. p. 22-6.

20. *Staggers N*, *Jennings BM*, *Lasome CE*. A usability assessment of AHLTA in ambulatory clinics at a military medical center. Mil Med 2010; 175: 518-24.

21. *Saitwal H*, *Feng X*, *Walji M*, *Patel V*, *Zhang J*. Assessing performance of an Electronic Health Record (EHR) using Cognitive Task Analysis. Int J Med Inform 2010; 79: 501-6.

22. *Gans D*, *Kralewski J*, *Hammons T*, *Dowd B*. Medical groups' adoption of electronic health records and information systems. Health Aff (Millwood) 2005; 24: 1323-33.

23. *Magrabi F*, *Ong MS*, *Runciman W*, *Coiera E*. An analysis of computer-related patient safety incidents to inform the development of a classification. J Am Med Inform Assoc 2010; 17: 663-70.

24. *Beuscart-Zephir MC*, *Anceaux F*, *Crinquette V*, *Renard JM*. Integrating users' activity modeling in the design and assessment of hospital electronic patient records: the example of anesthesia. Int J Med Inform 2001; 64: 157-71.